

# DATA TRIANGULATION IN ACTION: USING COMMENT ANALYSIS TO REFINE WEB QUALITY METRICS

Stuart J. Barnes, School of Management, University of East Anglia, Norwich NR4 7TJ, United Kingdom, stuart.barnes@uea.ac.uk

Richard T. Vidgen, School of Management, University of Bath, Bath BA2 7AY, United Kingdom, mnsrtv@bath.ac.uk

## Abstract

*This paper focuses on the evaluation of quality perceptions from users of an electronic government website. As government organizations have begun to enhance transparency, communicate and interact with citizens via the Web, developing appropriate online services has demanded heightened understanding of the requirements of users and appropriate tailoring of solutions. The site examined in the paper enables the online submission of self-assessed tax returns in the UK. Survey data collected via the eQual instrument were used to statistically model the perceptions of site users. However, in addition to quantitative data, we also collected open comments from respondents, and it is these comments that provide the crux of this paper. Such comments, via data triangulation, provided much more insight into the perceptions of site quality than the statistical data alone. The results of the comment analysis both support the instrument and point to additional factors determining the perceptions of quality of such e-government services requiring attention in the instrument development. The paper rounds off with a conclusions and an agenda for future research in this area.*

*Keywords: Online Tax Submission, Web Site, Evaluation, Quality, eQual, Data Triangulation, Comment Analysis.*

## 1 INTRODUCTION

Since the late-1990s, substantial government services have been provided via the Web in many countries such as Canada, the US, the UK, New Zealand, Australia, Portugal, Italy, Malaysia and Singapore. Broadly speaking, electronic government – defined as “the use of information technology, in particular the internet, to deliver public services in a much more convenient, customer-oriented, cost-effective, and altogether different and better way” (Holmes 2001) - has tremendous potential benefits. Indeed, 78 per cent of decision makers in government view e-government as having a positive effect on the business of government (Holden et al. 2003). The Bush administration, for example, has made e-government one of the five aspects of its core management agenda, citing the potential of e-government to enhance the efficiency and effectiveness of Federal agencies.

Government overarches every sector of society, and not only provides the legal, political, and economic infrastructure to support other sectors, but also exerts significant influence on the social factors that add to their development (Elmagarmid and McIver 2001). Electronic government thus spans many sectors and facets and has the potential to profoundly transform people’s perceptions of civil and political interactions with their governments. Through the Web, expectations of the service levels that e-government sites must provide have been raised considerably (Kubicek and Hagen 2001).

This research utilizes the eQual approach (previously called WebQual) to analyse user perceptions of the quality of a specific national Web site provided by the UK Government. eQual was developed originally as an instrument for assessing user perceptions of the quality of e-commerce Web sites. The

instrument has been under development since the early part of 1998 and has evolved via a process of iterative refinement in different e-commerce and e-government domains (e.g. see Barnes and Vidgen 2002, 2003a, 2003b, 2003c, 2004). The essence of the method focuses on turning qualitative customer assessments into quantitative metrics that are useful for management decision-making.

In the eQual method, metrics may be supplemented by open comments from respondents. If a large enough sample of these is provided this may allow a degree of qualitative triangulation and help to understand some of the “why” questions underlying statistical variance explanations. In this paper, the high number of responses and mix of qualitative and quantitative data allowed just that. Moreover, via a detailed comment analysis, we attempt to provide a detailed critique and refinement of the eQual instrument.

The Web site examined in this research is that of the Inland Revenue – a site relating to UK tax policy and administration. From an e-government perspective, this is a site that goes beyond information provision to interaction and transaction with citizens, the next major phase of e-government rollout being pursued by governments around the globe (Cohen and Eimicke 2003). As such, it touches on many aspects of e-government web quality that have much broader implications outside of the UK case, particularly for other governments following similar development paths.

The structure of the paper is as follows. In the next two sections we describe the background to the research and the methodology used. Sections four and five respectively report on the quantitative and qualitative data analyses. Conclusions are drawn in the last section.

## **2 BACKGROUND TO THE RESEARCH PROJECT**

A project to evaluate the quality of the UK Inland Revenue Web site (<http://www.ir.gov.uk/>) was initiated by the Tax Management Research Network, a consortium of tax practitioners and academics, in the early part of 2001.

Aside from information provision, a major part of the Inland Revenue’s Web site is the launch of a self-assessment facility for tax returns, first used for the 1999 to 2000 financial year to submit returns by 5 April 2001. Thus, the site provides a high degree of interactivity and the possibility for transactions. The online self-assessment facility is a major part of the Inland Revenue’s £200 million e-strategy (HMSO 2001) aimed at delivering fifty per cent of services electronically by 31 December 2002. In addition, the long-term aims are to provide all services electronically by 31 December 2005, by which time the take up of services should be 50 per cent. The proposed benefits for taxpayers of using the Self Assessment service are accuracy, convenience, confirmation of submission, and faster processing of any tax refunds (HMSO 2001). Whilst it is difficult to predict confidently the savings achievable, the department estimates that when take up reaches 50 per cent across all activities, this might enable efficiency saving equivalent to some 1,300 posts.

The evaluation of the IR Web site was undertaken using the eQual instrument, developed at the University of Bath, and was carried out during the period 1 August through 30 September 2001. In this report we present the results of analysis of the comments that subjects posted while completing the eQual survey. The standard quantitative eQual analysis is thus supplemented by qualitative comments of the respondents to provide triangulation of the results and a deeper insight into user attitudes (Barnes and Vidgen 2003b, 2004).

## **3 RESEARCH METHODOLOGY**

This section describes the quantitative and qualitative aspects of the research method and how they were operationalised. In particular, we describe the eQual instrument, the design of the evaluation and the methodology used for the qualitative comment analysis.

### 3.1 The eQual instrument

A review of the literature on Web site evaluation revealed no comprehensive instruments aimed specifically at e-government Web services. Therefore, and at the request of the Inland Revenue, we adopted the eQual instrument, adapting the format for interactive and non-interactive users. By adapting a previously developed and validated instrument, benefits accrue in the form of improved validity, the ability to compare results from previous studies with the current study and a movement towards building a cumulative tradition of research (Straub and Carlson 1989, Malhotra and Grover 1998).

eQual is based on quality function deployment (QFD), which is a “structured and disciplined process that provides a means to identify and carry the voice of the customer through each stage of product and or service development and implementation” (Slabey 1990). Applications of QFD start with capturing the ‘voice of the customer’ - the articulation of quality requirements using words that are meaningful to the customer. These qualities are then fed back to customers and form the basis of an evaluation of the quality of a product or service. eQual differs from studies that emphasise site characteristics or features (Kim and Eom 2002), which are used as part of later processes in QFD. In the context of eQual, Web site users are asked to rate target sites against each of a range of qualities and to rate each of the qualities for importance. Although the qualities in eQual are designed to be subjective, there is a significant amount of data analysis using quantitative techniques, for example, to conduct tests of the reliability of the eQual instrument.

Category	Questions
<i>Usability</i>	1. I find the site easy to learn to operate 2. My interaction with the site is clear and understandable 3. I find the site easy to navigate 4. I find the site easy to use 5. The site has an attractive appearance 6. The design is appropriate to the type of site 7. The site conveys a sense of competency 8. The site creates a positive experience for me
<i>Information Quality</i>	9. Provides accurate information 10. Provides believable information 11. Provides timely information 12. Provides relevant information 13. Provides easy to understand information 14. Provides information at the right level of detail 15. Presents the information in an appropriate format
<i>Service Interaction</i>	16. Has a good reputation 17. It feels safe to complete transactions 18. My personal information feels secure 19. Creates a sense of personalization 20. Conveys a sense of community 21. Makes it easy to communicate with the organization 22. I feel confident that goods/services will be delivered as promised
<b>OVERALL</b>	23 Overall view of the Web site

Table 1: The eQual Questionnaire

eQual has been under development since 1998 and has undergone numerous iterations. The development of eQual is discussed fully elsewhere (see Barnes and Vidgen 2002). eQual 4.0, as shown in Table 1, draws on research from three core areas:

- *Information quality* from mainstream IS research. A core part of the eQual instrument, from version 1.0, was the quality of online information. The questions developed in this segment of eQual build on literature focused on information, data and system quality, including Bailey and Pearson (1983), Strong et al. (1997) and Wang (1998).
- *Interaction and service quality* from marketing, e-commerce and IS service quality research. Bitner (1990, p. 72) adopts Shostack’s (1985) definition of a service encounter as “a period of

time during which a consumer directly interacts with a *service*” and note that these interactions need not be interpersonal - a service encounter can occur without a human interaction element. Bitner (1990) also recognizes that “many times that interaction *is* the service from the customer’s point of view” (p. 71). We suggest that interaction quality is equally important to the success of e-businesses as it is to “bricks and mortar” organizations (and possibly more so given the removal of the interpersonal dimension). In version 2.0 of the instrument we therefore extended the interaction aspects by adapting and applying the work on service quality, chiefly SERVQUAL (Parasuraman 1995, Zeithaml et al. 1990, 1993) and IS SERVQUAL (Pitt et al. 1995, 1997, Kettinger and Lee 1997, Van Dyke et al. 1997).

- *Usability* from human-computer interaction. In eQual 4.0 the usability dimension draws from literature in the field of human computer interaction (Davis 1989, 1993, Nielsen 1993) and more latterly Web usability (Nielsen 1999, 2000, Spool et al. 1999). Usability is concerned with the pragmatics of how a user perceives and interacts with a Web site: is it easy to navigate? Is the design appropriate to the type of site? It is not, in the first instance, concerned with design principles such as the use of frames or the percentage of white space, although these are concerns for the Web site designer who is charged with improving usability.

Notwithstanding, we have used quality workshops at every stage of eQual’s development to ensure that the qualities were relevant, particularly where they relate to pre-Internet literature and new organisational or industrial settings, such as e-government.

### 3.2 Design of the evaluation

The standard eQual instrument, previously called WebQual, contains 23 questions (Barnes and Vidgen 2002). These are shown in Table 1. Three of the questions relate to personal information and making transactions:

- Question 17: It feels safe to complete transactions
- Question 18: My personal information feels secure
- Question 22: I feel confident that goods/services will be delivered as promised

These three questions are relevant to respondents using the self-assessment facilities of the IR Web site but not to those who are using the site for information gathering purposes only. By self-assessment, we are referring to the online submission of tax returns that have been processed by the taxpayer using the self-assessment guidelines. The interaction questions were qualified with the instruction to “please tick n/a if you have not used the Internet service for self-assessment or the Internet service for PAYE”. This allows the data set to be divided between “information gatherers” and “interactors”.

The survey of Web site quality for the IR was conducted using an Internet-based questionnaire. The home page of the questionnaire had instructions and guidelines for completion of the instrument. From the home page the user opens a separate window (control panel) containing the Web site qualities to be assessed. The control panel allows the user to switch the contents of the target window between the instruction page, the IR Web site, and the quality dictionary. The online quality dictionary is linked to the question number, allowing the respondent to get a definition for any particular quality. Users were asked to rate the IR site for each quality using a scale ranging from 1 (strongly disagree) to 7 (strongly agree). Users are also asked to rate the importance of the quality to them, again using a 1 (least important) to 7 (most important) scale. Open comments were encouraged and a remarkably high proportion of respondents took the effort to provide an additional comment on the site (65%). These formed the basis of the comment analysis described in the next section.

The evaluation resulted in 420 usable responses. In line with previous studies using eQual 4.0, demographic and other respondent information were also collected from the sample. In particular, we were interested in the age, sex and type of user, their use of the site, and their experience and use of the Internet in general. The respondents were typically highly experienced and intensive users of the Internet, although not intensive users of the IR Web site. The majority of respondents were male

(71%) and of a working age. 10% use the IR site daily. Agents and accountants comprised 15.5% of respondents, while 60% categorized themselves as “other”.

### **3.3 Qualitative analysis of open comments**

Alongside the quantitative analysis, the open comments of respondents were used to perform a detailed qualitative analysis. In particular, we aimed to summarize the themes regarding user perceived website qualities which were mentioned in the open comments, compare these themes with the eQual questions, and perform a degree of triangulation with the quantitative analysis.

Comments were imported into NVivo (a tool for qualitative analysis) as documents with demographic data of the respondents as attributes. NVivo was then used to code the data. In order to compare and contrast this data with the eQual framework, the codes regarding user perceived website qualities were not extracted from the eQual questionnaire, but emerged initially from the preliminary coding process and then were refined through several coding iterations. No structure was imposed during the preliminary coding. Every sentence was covered at least by one code. The second and third rounds of coding were based around comparatively examining the text under the same codes through the view provided by NVivo. The tree structure was re-organized according to the examination. The purpose of the third and final round of coding was particularly to ensure consistency and accuracy. When the coding process had finished completely, these codes were categorized according to the dimensions of the eQual framework.

In addition to the codes associated with assessing web quality, the attitude expressed via each comment was also made explicit by codes such as “criticism”, “praise”, “suggestion” and “untitled” (which means it is not a relevant comment). The purpose of this code set is to help triangulate the quality scores given by the respondents. For this set of codes, the unit of coding is the whole piece of comment. Every comment was covered at least by one code.

## **4 QUANTITATIVE RESULTS**

For the quantitative analysis, we were particularly interested to discover the aspects of the eQual instrument that determine the user’s overall perception of the quality of the IR site. We were also interested in the quality priorities of each user, indicated by a measure of importance for each question. The data analysis below was conducted on the weighted dataset, where the rating for a question for each respondent is multiplied by its perceived importance. A summary of the results is presented here for comparative purposes; for full details see Barnes and Vidgen (2003b, 2004).

One key aim of this approach is to achieve some overall quality rating for the Web site so that we can benchmark the perceptions of site users. The total scores make it difficult to give a standard benchmark for the Web site, especially since questions 17, 18 and 22 are omitted from the responses of non-interactive users. One way to achieve this is to index the total weighted score for a site against the total possible score (i.e. the total importance for all questions answered multiplied by 7, the maximum rating for a site). The result is expressed as a percentage. A summary of these calculations and totals are given in Table 2.

Overall, we can see quite clearly that the interactive users benchmarked well below the non-interactive users (62% and 72% respectively), a difference of 10 points in the eQual Index (EQI). Even more remarkable is that the evaluations of interactive users rated consistently below that of non-interactive users for all questions, with differences ranging from 1 to 18 points. The largest differences relate to usability (items 1, 4, 2, 3), followed by competency and understandable information.

The data indicates differences in perceptions in terms of eQual site quality. Here we examine where these perceived differences have occurred and consider the overall shape of the evaluation of the IR site. Previous research for eQual has led to a number of valid and reliable question subgroupings (see

section 3). These categories provide some useful criteria by which to assess the perceptions of site users.

No.	Description	Max. Score (I)	Interact		Max. Score (NI)	No Interaction		Difference (EQI2 - EQI1)
			Wgt. Score	EQI1		Wgt. Score	EQI2	
1	I find the site easy to learn to operate	42.14	23.41	56%	42.80	31.46	74%	-18%
2	My interaction with the site is clear and understandable	41.92	23.61	56%	41.95	30.08	72%	-15%
3	I find the site easy to navigate	42.51	23.77	56%	43.82	30.48	70%	-14%
4	I find the site easy to use	43.06	23.89	55%	43.91	31.41	72%	-16%
5	The site has an attractive appearance	30.43	19.56	64%	28.52	19.79	69%	-5%
6	The design is appropriate to the type of site	33.12	22.65	68%	33.86	26.55	78%	-10%
7	The site conveys a sense of competency	39.42	24.77	63%	38.57	29.79	77%	-14%
8	The site creates a positive experience for me	36.98	18.87	51%	34.95	22.08	63%	-12%
9	Provides accurate information	44.50	31.89	72%	45.95	37.76	82%	-11%
10	Provides believable information	43.30	32.90	76%	46.27	39.74	86%	-10%
11	Provides timely information	41.94	29.23	70%	45.38	35.08	77%	-8%
12	Provides relevant information	42.79	30.24	71%	45.88	35.86	78%	-7%
13	Provides easy to understand information	42.91	25.01	58%	44.11	31.81	72%	-14%
14	Provides information at the right level of detail	41.08	25.42	62%	43.33	29.65	68%	-7%
15	Presents the information in an appropriate format	39.84	26.39	66%	40.71	30.45	75%	-9%
16	Has a good reputation	36.69	22.36	61%	37.35	27.43	73%	-12%
17	It feels safe to complete transactions	42.26	30.67	73%	44.92	-	-	-
18	My personal information feels secure	42.81	31.90	75%	45.18	-	-	-
19	Creates a sense of personalization	31.84	16.00	50%	25.91	13.31	51%	-1%
20	Conveys a sense of community	26.12	12.25	47%	21.06	10.30	49%	-2%
21	Makes it easy to communicate with the organization	39.12	19.80	51%	37.49	20.54	55%	-4%
22	I feel confident that goods/services will be delivered as promised	41.00	23.23	57%	43.56	-	-	-
TOTALS:		865.76	537.81	62%	875.49	533.59	72%	-10%

Note: n=420; interactive users = 264; non-interactive users = 156

Table 2. Weighted scores and eQual indices – interactive and non-interactive users

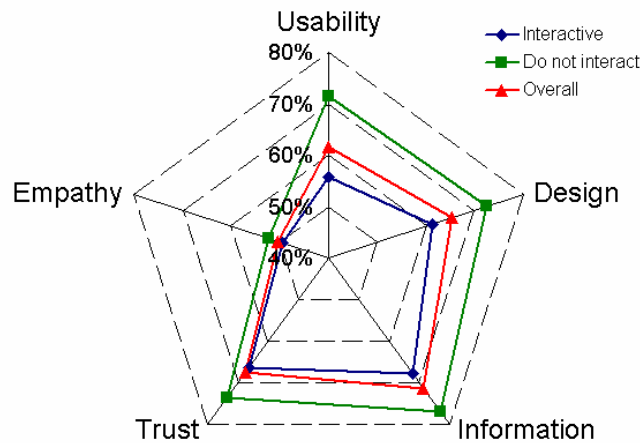


Figure 1. Radar chart of eQual subcategories for user groups

As a starting point, the data was summarised around the questionnaire subcategories. Then, and similarly to the eQual Index in Table 2, the total score for each category was indexed against the maximum score (based on the importance ratings for questions multiplied by 7). Figure 1 is the result, which rates the two sets of users with these criteria. Note that the trust category is limited to question 16 for the users who 'do not interact'. Further, the scale has been adjusted to between 40% and 80% to allow for clearer comparison. Clearly the users who do not interact with the site have higher perceptions in all aspects, although the general pattern of site ratings is similar for all users.

In absolute terms, for users who ‘do not interact’ all site categories rate quite highly at between 72% and 77%, except for empathy (52%). Although this category also rates lowest in importance, it does indicate an opportunity for building relationships with users. For ‘interactive’ users, empathy, usability and design rate lowest (at 49%, 56% and 61% respectively), with information (68%) and trust (66%) the best rated scores.

Figure 1 demonstrates that the biggest subcategory differences in perceptions are in usability and design – 16% and 11% respectively. Close behind is information quality – at 9%. The most similar quality perceptions were for empathy – a difference of just 3%. Apparently, interaction with the Inland Revenue site severely affects perceptions of usability and design.

## **5 QUALITATIVE RESULTS**

In this section we present the results of the qualitative analysis of respondent comments. These are further analysed and compared to the quantitative data in the next section.

### **5.1 Sampling issues**

In order to triangulate the quantitative analysis on the whole sample (420 cases) with the findings from the comments of 273 respondents, it is necessary to verify that these 273 cases are a random selection. These were tested using ANOVA. The compared statistics include the proportion of interactive and non-interactive users, the distribution of demographics and web experiences of the respondents, and the means, standard deviations and standard errors of the means of the 23 items of the two samples. The result of the comparison showed that the sample is a random selection from the 420 cases.

### **5.2 Results of the comment analysis**

Table 3 lists the themes covered by the coded open comments and their occurrence frequencies. The code occurrence frequencies, which are assumed to be a measure of the relevance and importance of the themes to the respondents, were calculated in NVivo. The statistical mean of the code occurrence frequency is 14.5, the median is 12, and the upper quartile is 23.25.

The coded comments in Table 3 have not yet been organised into any type of common groupings. Therefore, although we can see clear areas of common interest among the respondents’ comments, this makes systematic comparison with the quantitative data difficult. To facilitate triangulation, the eQual framework was used to organise the codes into four groupings: three conceptual groups from the eQual instrument, plus an ‘Other’ group where placement of a specific comment code proved difficult. The results are shown in Table 4.

The next challenge is to make some assessment of the importance and rating of qualities, and thus quality categories, as inferred from the comment analysis data. It is reasonable to assume that those areas commented on most by respondents were also considered the most important. This is particularly so considering that the respondents typically only mentioned a few topics of importance in the short statements that they added to the end of the eQual questionnaire; effort is likely to be focused on topics of immediate concern.

The next challenge is to make some assessment of the importance and rating of qualities, and thus quality categories, as inferred from the comment analysis data. It is reasonable to assume that those areas commented on most by respondents were also considered the most important. This is particularly so considering that the respondents typically only mentioned a few topics of importance in the short statements that they added to the end of the eQual questionnaire; effort is likely to be focused on topics of immediate concern.

Code	Occurrence Frequency	Attitude			
		criticism	praise	suggestion	untitled
Navigation	43	36	6	1	0
Locating information	41	36	5	0	0
Information provision	33	18	1	11	3
Form completion	31	28	1	2	0
Usefulness	30	7	22	0	1
Search facility	25	24	0	1	0
Communication with organisation	25	17	1	7	0
Ease of use	18	14	4	0	0
How informative	17	3	14	0	0
Authentication	17	16	1	0	0
Organization and format	13	10	3	0	0
Responsiveness	12	11	1	0	0
Online help	12	7	0	5	0
Experience with the site	12	9	3	0	0
Function provision	12	4	0	8	0
User-friendliness	11	9	2	0	0
Look and feel	11	5	5	1	0
Clarity	10	7	3	0	0
System performance	7	5	2	0	0
Currency	7	3	1	3	0
Instruction	7	6	0	1	0
Notification	5	3	0	2	0
Simplicity	4	4	0	0	0
Consistency	3	3	0	0	0
Accuracy	1	1	0	0	0
Security	1	0	1	0	0
Feedback mechanism	1	1	0	0	0
Personalization need	1	0	0	1	0
Advertisement	1	0	0	0	1

Table 3. The coding scheme and code occurrence frequency

Category	Comment Code	Occurrence Frequency	eQual Equivalence
<i>Usability</i>	Navigation	43	Navigation (Q3)
	Locating information	41	Navigation (Q3)
	Form completion	31	Usability (Q1, 2, 4)
	Search facility	25	* not a quality
	Ease of use	18	Usability (Q1, 2, 4)
	Experience with the site	12	Site design (Q8)
	User-friendliness	11	Usability (Q1, 2, 4)
	Look and feel	11	Site design (Q5)
	System performance	7	** no equivalent quality
	Instruction	7	Usability (Q1-4)
	Simplicity	4	Site design (Q6)
<i>Information</i>	Information provision	33	General (Q9-15)
	Usefulness	30	Relevance (Q12)
	How informative	17	General (Q9-15)
	Organization and format	13	Format (Q15)
	Clarity	10	Easy to understand (Q13)
	Currency	7	Timeliness (Q11)
	Consistency	3	General (Q9-15)
	Accuracy	1	Accuracy (Q9)
<i>Service</i>	Comm. with organisation	25	Communication (Q21)
	Authentication	17	Security (Q18)
	Responsiveness	12	Communication (Q21)
	Online help	12	Communication (Q21)
	Notification	5	Communication (Q21)
	Security	1	Security (Q18)
	Feedback mechanism	1	Communication (Q21)
	Personalization need	1	Personalisation (Q19)
<i>Other</i>	Function provision	12	* not a quality
	Advertisement	1	* not a quality

Table 4. Comment codes organized according to the eQual framework

In order to gauge the level of importance, it is useful to apply some rudimentary statistics to the data, as shown in Table 5. First, comments were grouped into baskets, according to eQual subcategories:

usability; site design; information; trust and empathy. This provided some absolute numbers on numbers of comments. Next, an average expected number of comments was calculated for each category; this was based on the average number of comments for each comment code multiplied by the number of comment codes in each category. Table 5 shows the difference between these in absolute terms as well as the proportion (as a percentage) of actual to expected comments. Based on the results, it appears that Usability is considered to be the most important category (rated 153% of expected comments), whilst Information Quality is rated as moderate importance (97% of expected). All the other categories are rated as low importance, with between 61% and 63% of expected comments.

eQual Category	Coded Comments	Actual	Ave. Expected	Difference	Actual/Expected	Importance
<i>Usability</i>	# 1-3, 5, 7, 9-10	158	103	55	153%	High
<i>Site design</i>	# 6, 8, 11	27	44	-17	61%	Low
<i>Information</i>	# 12-19	114	118	-4	97%	Moderate
<i>Trust</i>	# 21, 25	18	29	-11	61%	Low
<i>Empathy</i>	# 20, 22-24, 26-27	56	88	-32	63%	Low

Table 5. Assessing the importance of comment categories

The next step of the assessment is to make some judgement of the perceived quality of the Inland Revenue site as indicated by the respondents' comments. For this, we focus on those comments that have been interpreted to make a value judgement (either positive or negative) – other comments are ignored. Table 6 lists the numbers of positive, negative and combined comments in the eQual categories. Using these totals, we can make some analysis of the social consensus regarding perceived qualities. Two columns provide some indication of this: first, the proportion of positive comments from the total; and second, the ratio of positive to negative comments. As we can see, each of these assessments falls well below that of the quantitative analysis above. However, the pattern is very similar.

eQual Category	Comments	(-)ve	(+)ve	(+)ve + (-)ve	(+)ve/Total	(+)ve/(-)ve	Rating*(A/E)
<i>Usability</i>	# 1-3, 5, 7, 9-10	134	20	154	13%	15%	20%
<i>Site design</i>	# 6, 8, 11	18	8	26	31%	44%	19%
<i>Information</i>	# 12-19	52	44	96	46%	85%	44%
<i>Trust</i>	# 21, 25	16	2	18	11%	13%	7%
<i>Empathy</i>	# 20, 22-24, 26-27	39	2	41	5%	5%	3%

Table 6. Assessing the perceived quality of comment categories

In order to provide a weighted score, based on both perceptions of importance of qualities and value judgements about the perceived quality of the Inland Revenue site, a radar chart was created. This is based on the actual/expected comments from the total. The results are easily comparable to the quantitative analysis above. The scale is from 0% to 50% (Figure 1); the qualitative analysis is more similar to that of the quantitative analysis (Figure 2). Note that the pattern is very similar to that of the quantitative analysis.

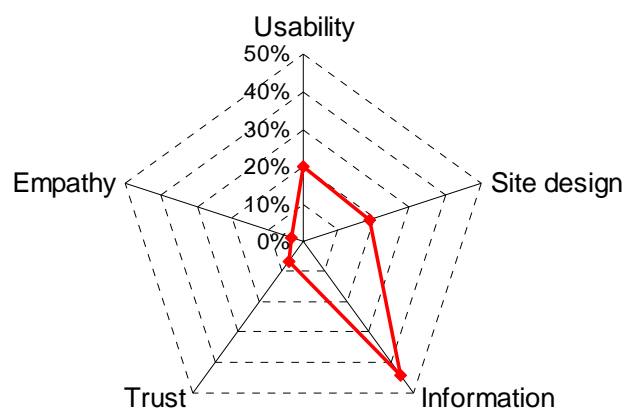


Figure 2. Radar chart of eQual subcategories based on comment analysis

### 5.3 Discussion

The qualitative results appear to provide some useful evidence and support the quantitative results. For example, we can see that Empathy has the lowest weighted score, while Information Quality has the highest. The other areas of quality fall in between. The most marked difference is the low Trust score, close to Empathy, but which was more similar to Usability and Site Design in the quantitative results.

While the pattern is similar, the basic mode of data collection and analysis appears to provide indicative information only, i.e., while the pattern is similar, the level of numerical assessment is quite different. One explanation for this is that those proportion of individuals who made comments – on which the qualitative analysis relies – had specific points to make, and that most of those points were critical, skewing the assessment downwards. The quantitative data is likely to be broader, taking onboard another 147 individuals who did not make comments and whose views may be quietly neutral or positive. This is likely to have a smoothing effect on the evaluations, providing a degree of moderation.

Overall, the comments collected tally well with the eQual instrument (Table 5). Some comments tally with specific qualities, while others are more general or provide specific instances of a quality. Four comment areas do not easily match the eQual instrument. Three of these (#4, #28 and #29) are not actually qualities in the true sense (Slabey). Rather, they refer to parts – either specifically or generally. However, one of the comment areas (#9) is very relevant, and does not have adequate coverage in the eQual instrument. We therefore intend to include an item for this quality in the future.

## 6 SUMMARY AND CONCLUSIONS

This research has examined an important area of development for digital government – online taxation systems. It focuses on the experiences in the UK surrounding the introduction of an online facility for self-assessed tax returns, and specifically, in evaluating the factors impacting on user perceptions. eQual is based on user perceptions of quality weighted by importance. Within eQual, five factors are used: usability, design, information, trust, and empathy. In this study we have evaluated these using both qualitative and quantitative methods to provide a degree of triangulation.

The results demonstrate that the use of comment analysis alongside traditional survey data can provide a very useful method of triangulation, adding strength to the results of the web quality assessment. The novel method of quantifying respondent comments used in this paper provides a contribution to data triangulation for web quality assessment. In the quantitative results, interaction was a clear determinant of the user's perception of overall website quality. In a previous paper, we found a distinct and consistently different rating of the site between two user groups: information seekers and interactors (Barnes and Vidgen 2004). The latter group involves those who attempted to engage in online self-assessed tax returns, and who typically rated the quality of the site much lower than those who merely sought information. Key problems affecting the perceptions of the interactive users are the usability of the self-assessment facility and difficulty communicating with the organisation. The qualitative comment analysis underlines this and the comments made by respondents (which totalled 273 out of 420) were typically critical and from the interactor camp.

The research findings suggest that while information quality is perceived well – both in qualitative and quantitative results – respondents are more critical of site design and usability, particularly in the comment analysis. This finding is also borne out in indicative information about submission experiences. System logs showed that nearly four out of five attempted submissions in the 1999 to 2000 round did not succeed first time. The proportion of successful attempts for first time submission had only reached 44 per cent on average between April and September 2001, and it improved further to an average of 70 per cent for the quarter ending December 2001 (HMSO 2002).

Another major finding is the low perceived level of trust and empathy. The perceived evaluation of trust is perhaps not surprising given the ‘necessary evil’ that tax affairs in general and the IR in particular are usually associated with. However, the need for empathy (particularly communication) in the delivery of services is an interesting finding. Recent developments at the Inland Revenue also support these findings; the Inland Revenue is currently moving from its existing arrangements for taxpayers to file a tax return towards a ‘portal’ environment offering secure personalised services, such as the option for taxpayers to view their account, communicate with the Inland Revenue, as well as the facility to file a tax return electronically.

Overall data triangulation has provided fruitful results for the research, including areas of web assessment that are currently underrepresented in the eQual instrument. This has integrated new ideas for quantifying qualitative data (perceptions, importance) for use in web quality assessment activities. Future research will aim to further develop and refine the method, the instrument and integrate reflective-learning about e-government services. We also aim to learn more about the domain specific qualities of e-government and build a clearer picture of the perceptions of quality for the wide range of new services that are emerging in this area.

## **7 ACKNOWLEDGEMENT**

We wish to acknowledge the help of Xiaofeng Wang, a Master’s student at the University of Bath, who provided the initial coding of the comments used in this analysis.

## **References**

- Bailey, J.E. and Pearson, S.W. (1983). Development of a tool for measuring and analyzing computer user satisfaction. *Management Science*, 29 (5), 530-44.
- Barnes, S.J. and Vidgen, R.T. (2002). An integrative approach to the assessment of e-commerce quality. *Journal of Electronic Commerce Research*, 3 (3), 114-127.
- Barnes, S.J. and Vidgen, R.T. (2003a). Measuring Web site quality improvements: a case study of the Forum on Strategic Management Knowledge Exchange. *Industrial Management and Data Systems*, 103 (5), 297-309.
- Barnes, S.J. and Vidgen, R.T. (2003b). Interactive E-Government: Evaluating the Web Site of the UK Inland Revenue. *Journal of Electronic Commerce in Organizations*, 2(1), 22pp.
- Barnes, S.J. and Vidgen, R.T. (2003c). Evaluating alcohol advice Web sites: a cross-country perspective. *International Journal of Management Literature*, 2 (2), pp. 51-64.
- Barnes, S.J. and Vidgen, R.T. (2004). Interactive eGovernment Services: Modelling User Perceptions with eQual. *Electronic Government*, 1(2), 213-228.
- Bitner, M. (1990). Evaluating service encounters: the effects of physical surroundings and employee responses. *Journal of Marketing*, 54, 69-82.
- Cohen, S. and Eimicke, W. (2003). The future of e-government: a project of potential trends and issues. In *Proceedings of the 36th Hawaii International Conference on System Sciences*, Big Island, Hawaii, January.
- Davis, F. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13 (3), 340-51.
- Davis, F. (1993). User acceptance of information technology: System characteristics, user perceptions, and behavioral impacts. *International Journal of Man-Machine Studies*, 38, 475-487.
- Elmagarmid, A. K., and McIver, W. J. (2001). The ongoing march toward digital government. *IEEE Computer*, 34 (2), 32-8.
- HMSO (2001). *Inland Revenue e-Strategy*. HMSO, London.
- HMSO (2002). *e-Revenue*. HMSO, London.

- Holden, S.H., Norris, D.F. and Fletcher, P.D. (2003). Electronic government at the grass roots: contemporary evidence and future trends. In Proceedings of the 36th Hawaii International Conference on Systems Sciences, Big Island, Hawaii, January.
- Holmes, D. (2001). *E.Gov: E-Business Strategies for Government*. Nicholas Brealey, London.
- Kettinger, W. and Lee, C. (1997). Pragmatic perspectives on the measurement of information systems service quality. *MIS Quarterly*, 21, 223-240.
- Kim, E.B. and Eom, S.B. (2002). Designing effective cyber store user interface. *Industrial Management and Data Systems*, 102 (5), 241-51.
- Kubicek, H., and Hagen, M., (2001). Integrating e-commerce and e-government: the case of Bremen Online Services. In *Designing E-Government* (Prins, J. Ed.), Kluwer Law International, The Hague.
- Malhotra, M. and Grover, V. (1998). An assessment of survey research in POM: From construct to theory. *Journal of Operations Management*, 16 (4), 403-423.
- Nielsen, J. (1993). *Usability Engineering*. Morgan Kaufmann, San Francisco.
- Nielsen, J. (1999). User interface directions for the Web. *Communications of the ACM*, 42 (1), 65-72.
- Nielsen, J. (2000). *Designing Web Usability*. New Riders Publishing, Indiana.
- Parasuraman, A. (1995). Measuring and monitoring service quality. In *Understanding Services Management* (Glynn, W. and Barnes, J. Eds.), Wiley, Chichester.
- Pitt, L., Watson, R. and Kavan, C. (1997). Measuring information systems service quality: Concerns for a complete canvas. *MIS Quarterly*, 21, 209-221.
- Pitt, L., Watson, R., and Kavan, C. (1995). Service quality: a measure of information systems effectiveness. *MIS Quarterly*, 19 (2), 173-87.
- Shostack, G. (1985). Planning the service encounter. In *The Service Encounter* (Czepiel, J. Solomon, M. and Surprenant, C. Eds.), Lexington Books, Lexington, MA.
- Slabey, R. (1990). QFD: A basic primer. Excerpts from the implementation manual for the three day QFD workshop. In *Transactions from the Second Symposium on Quality Function Deployment*, Novi, Michigan, June 18-19.
- Spool, J., Scanlon, T., Schroeder, W., Snyder, C. and DeAngelo, T. (1999). *Web Site Usability: a Designer's Guide*. Morgan Kaufmann, San Francisco.
- Straub, D.W. and Carlson, C.L. (1989). Validating instruments in MIS research. *MIS Quarterly*, 13 (2), 147-169.
- Strong, D., Lee, Y. and Wang, R. (1997). Data quality in context. *Communications of the ACM*, 40 (5), 103-110.
- Van Dyke, T., Kappelman, L. and Prybutok, V. (1997). Measuring information systems service quality: Concerns on the use of the SERVQUAL questionnaire. *MIS Quarterly*, 21, 195-208.
- Wang, R.Y. (1998). A product perspective on Total Data Quality Management. *Communications of the ACM*, 41 (2), 58-65.
- Zeithaml, V. A., Berry, L. and Parasuraman, A. (1993). The nature and determinants of customer expectations of service. *Journal of the Academy of Marketing Science*, 21 (1), 1-12.
- Zeithaml, V.A., Parasuraman, A., and Berry, L. (1990). *Delivering Quality Service: Balancing Customer Perceptions and Expectations*. The Free Press, New York.