

International Journal of Geographical Information Science

ISSN: 1365-8816 (Print) 1362-3087 (Online) Journal homepage: www.tandfonline.com/journals/tgis20

Classification through consultation: public views of the geography of the e-Society

Paul A. Longley & Alexander D. Singleton

To cite this article: Paul A. Longley & Alexander D. Singleton (2009) Classification through consultation: public views of the geography of the e-Society, International Journal of Geographical Information Science, 23:6, 737-763, DOI: [10.1080/13658810701704652](https://doi.org/10.1080/13658810701704652)

To link to this article: <https://doi.org/10.1080/13658810701704652>



Published online: 05 Aug 2009.



Submit your article to this journal 



Article views: 331



View related articles 



Citing articles: 2 [View citing articles](#) 

Research Article

Classification through consultation: public views of the geography of the e-Society

PAUL A. LONGLEY* and ALEXANDER D. SINGLETON

Department of Geography, University College London, London, UK

(Received 30 July 2007; in final form 22 September 2007)

Although viewed by business and commerce as successful solutions, geodemographic profiling of neighbourhoods has attracted wide-ranging criticism in the academic literature. This paper addresses some specific concerns that arise because the derivation of classifications is rarely transparent and open to scrutiny or challenge. The substantive focus of the research reported in this paper is a nation-wide geodemographic classification of how people engage with new information and communication technologies (ICTs). In response to the critique of geodemographics as a 'black box' technology, we describe how the classification was opened up to public scrutiny and how we conducted a major consultation exercise into the reliability of its results. We assess the message of the 50,000+ searches and 3952 responses collected during the consultation exercise, in terms of possible systematic errors in the shape and detail of the classification. Unusually for Internet-based surveys, we also investigate the likely reliability of the response information received and identify ways in which the outcome of consultation might be used to improve the classification. We believe that this is the first-ever large-scale consultation survey of the validity and remit of a geodemographic classification and that it may have wider implications for the creation of geodemographic classifications.

Keywords: Geodemographics; Information and communications technologies; e-Society

1. Introduction

This paper reports on the results of a major public consultation exercise into the validity of a application-specific geodemographic classification, which had been created to present a picture of the engagement of Great Britain's citizens with new information and communication technologies (ICTs). We believe the exercise to be one of the largest geographically extensive and representative Internet-based surveys to have been undertaken in the realm of geodemographics – a field in which there is rarely the opportunity for people to 'appeal' against the neighbourhood class to which they have been assigned. The results are of added poignancy in the light of long-standing if somewhat contradictory concerns about the use of geodemographics to target the provision of goods and services for private or social gain. On the one hand, there are concerns that geodemographics is of sufficient power to rewrite entire landscapes of consumption, if their very success in improving the targeting activities of private sector organisations leads them to

*Corresponding author. Email: plongley@geog.ucl.ac.uk

focus upon their best prospects to the detriment of other groups in society (Goss 1995). (If sustained, however, this argument strengthens the case for use of geodemographic applications in social marketing: see Longley 2005). On the other hand, other commentators (e.g. Curry 1995) have instead emphasised from a conceptual standpoint the ultimate futility of all attempts to generalise about the behaviour of the public.

Although contradictory, these arguments are each lent some credence by the 'black box' nature of many geodemographic classifications, particularly those marketed by the private sector, and the finality with which secret and apparently mechanistic solutions are devised and applied. Response to these criticisms suggests not only a need for greater transparency and humility in building classifications but also a role for consultation in evaluating the outcomes of the classification process. If geodemographics is indeed a powerful tool of the surveillance society, soliciting and accommodating feedback might pinpoint residual errors in classification, while simultaneously making those that undertake classification more accountable to a broader constituency of interests. Conversely, if geodemographics is not in fact very powerful at all, feedback fulfils a much more important role with regard to the detection of inaccuracy, with the more immediate end of improving the overall outcome of the classification process. Such feedback might be gathered as part of a consultation exercise, involving the widest range of stakeholders that might be affected by use of any given classification. Ideally, then, the best classification schema can benefit from constructive feedback and consultation, in ways that make geodemographic analysis a more active and constructive partner in what has been termed 'social politics' (Burrows and Ellison 2004).

The remainder of this paper is structured as follows. In the next section, we describe the motivation and methodology underpinning the creation of the 'e-Society' classification, and the concerns that some users might have in applying it to analysis of ICT usage. We then describe how we conceived and implemented a consultation exercise as a way of evaluating the classification. We then evaluate the feedback that we received using checks internal to the classification and cross validation with external sources. In a concluding section, we discuss the usefulness of this consultation exercise and its contribution towards the development of geodemographic classification systems for other niche applications, especially in the public sector.

2. Background

The geodemographic classification that forms the basis to this study is a Great Britain-wide study of how people engage with new information and communications technologies. The research was funded as part of the Economic and Social Research Council 'e-Society' programme,¹ in part to provide countrywide context to other case studies in what remains the largest ever-academic research programme to investigate the adoption and use of new ICTs on society. A detailed description of the derivation of the classification is available in Longley *et al.* (2008) or at http://www.spatial-literacy.org/inc/resources/e_soc.pdf. Very briefly, the classification allocated every individual on the GB Electoral register to one of 23 Types and these in turn were aggregated into the eight Groups described in Table 1. Our motivation in creating this classification was to identify the spectrum of uses to which members of the public put ICTs and hence to describe levels of engagement with the 'e-Society'. Using this classification, it is possible to produce aggregate indicators of ICT usage for any convenient areal aggregation, and such aggregations are of obvious potential use in

Table 1. The e-Society classification.

e-Society Group	e-Society Group postcode frequency (%)	e-Society Type	e-Society Type postcode frequency (%)
A: E-unengaged	440,824 (23.5)	A01 Low technologists A02 Cable suffices A03 Technology as fantasy A04 Mobile's the limit A05 Too old to be bothered A06 Elderly marginalised	128,807 (6.9) 57,166 (3.0) 77,951 (4.2) 113,553 (6.1) 14,851 (0.8) 48,496 (2.6)
B: E-marginalised	114,098 (6.1)	B07 The Net; What's that? B08 Mobile explorers B09 Cable TV heartland	10,978 (0.6) 34,719 (1.9) 68,401 (3.6)
C: Becoming engaged	89,862 (4.8)	C10 E-bookers and communicators C11 Peer group adopters	46,176 (2.5) 43,686 (2.3)
D: E for entertainment and shopping	290,456 (15.5)	D12 Small time net shoppers D13 E for entertainment	183,282 (9.8) 107,174 (5.7)
E: E-independents	181,396 (9.7)	E14 Rational utilitarians E15 Committed learners E16 Light users	98,777 (5.3) 26,698 (1.4) 55,921 (3.0)
F: Instrumental e-users	161,825 (8.6)	F17 Computer magazine readers F18 E for financial management F19 Online apparel purchasers F20 E-exploring for fun	55,803 (3.0) 5561 (0.3) 60,380 (3.2) 40,081 (2.1)
G: E-business users	78,952 (4.2)	G21 Electronic orderers	78,952 (4.2)
H: E-experts	44,742 (2.4)	H22 E-committed H23 E-professionals	37,380 (2.0) 7,362 (0.4)

Unknown and Business Postcodes = 473,879 (25.3%).

exploring the geography of 'digital exclusion' or 'digital unengagement' in British society. Such information is the more important in view of the absence of questions (unlike some other countries) in the Census of Population on computer ownership or usage.

The classification was built using the 'Mosaic Pixel' classification method (Webber 2004a), and also utilised a range of private sector data sources supplied by Experian (Nottingham, UK: <http://www.business-strategies.co.uk>) as an 'in kind' contribution towards the research. The major advantage gained by utilising these data sources is the richness of the information that they provide about the respondents to private sector surveys and 'lifestyles' shopping questionnaires. A potential weakness of such sources is the voluntary nature of response to market research surveys and the unscientific approach to the collection of lifestyles data. Many private sector geodemographic classifications utilise 'pen portraits', comprising summary text and illustrative material, to summarise the prevailing characteristics of geodemographic groups and types, in order to render classifications more intelligible to users. We adopted this industry practice for purposes of our 'e-Society' classification, and the pen portraits may be viewed at <http://www.spatial-literacy.org/esocietyprofiler/eclassification.php>. A summary of the Group and Type labels in our classification is presented in Table 1, along with counts of the numbers of residential unit postcodes assigned to each using the procedure described in Section 3 below. One issue that is

immediately apparent from the classification is the high incidence of the 'E-unengaged' Group (dominant in 23.5% of all unit postcodes): this is, in practice, a 'catch all' category comprising individuals for whom we had little or no evidence of use of ICTs at the time the classification was built.

Over the last 30 years, general-purpose geodemographic classifications have become established as tactical and strategic marketing tools throughout the private sector, and today are used by almost all customer facing organisations (Sleight 2001; Harris *et al.* 2005). However, this success is very much predicated upon the perception that geodemographics allows improvement upon simple 'equal share' estimates of the demand for goods and services that are based upon undifferentiated counts of population size. There has been little published evaluation of the performance of geodemographics against equal share benchmarks (most probably because of the sensitive commercial nature of the classifications: but see Webber 2004b), and it is the case that geodemographics remains a 'black box' technology to most users. Issues of the provenance of data used to build classifications, and the weighting schemes applied to the variable selected for inclusion are not widely understood, and this has often restricted public sector applications where the interests and representation of the widest range of stakeholders must be acknowledged and accommodated.

In this context, we believe that a number of issues are germane to the wider adoption and use of geodemographics, particularly in public sector applications in which fairness and transparency are of utmost importance. Still other issues are important to ascertain whether bespoke geodemographics such as our 'e-Society' classification provide better ways of understanding niche applications than the general-purpose geodemographics of the private sector or general-purpose indicators (such as the 2004 Index of Multiple Deprivation:<http://www.communities.gov.uk/archived/publications/communities/indicesdeprivation>) that are presently used in the public sector.

3. Evaluating the 'e-Society' classification

In conventional academic terms of peer review our 'e-Society' project was a success and indeed was one of just two projects in the first phase of the initiative to be evaluated by ESRC as 'outstanding' in its research achievements. At the suggestion of the ESRC Programme Director (Burrows, personal communication) and following sustained public interest in the results of one of our other research projects (the origins of Anglo-Saxon surnames: see <http://www.spatial-literacy.org/UCLnames/>), we repackaged the classification in a form suitable for scrutiny by members of the public.

Our intention was to solicit feedback to our 'e-Society' classification, to identify the possible deficiencies in it, and to break down the responses in order to try to validate the classification that we had devised. The feedback was solicited through the medium of the Internet and, cognizant that Internet surveys are notoriously vulnerable to selection and response bias (Jones 1999; De Vaus 2002), we undertook both internal and external validation of the results: internal validation of the results in terms of the searches that were made within the classification and user opinions of its characteristics and external validation of suggested improvements to the classification with reference to a general-purpose geodemographic classification.

The most widely recognised level of aggregation for implementing our consultation exercise—which also involved discarding the least information from the individual-level classification—was the unit postcode. The most prevalent e-Society

Type at the level of the individual citizen was assigned to each domestic unit postcode. We devised an easily navigable user interface that was similar in design to that used in the publicity media produced by commercial purveyors of geodemographics (see Figure 1 and <http://www.spatial-literacy.org/esocietyprofiler/>), with hotlinks to an overview of the project, to the pen portraits of the classification, and to the feedback area of the site. Upon entering a valid residential unit postcode, users are presented with a location map (obtained from Google Maps), a description of the Group and Type that has been assigned to the unit postcode, and a hot-linked list of the 10 most similar postcode districts in terms of unit postcode composition. Error messages were used to flag invalid postcodes, such as those pertaining to non-residential buildings, or those postcodes which post-dated the creation of the utility. Additionally, a prominent link was provided to allow users to report perceived errors in the classification through a drop down menu of Groups and Types.

The previous experience of our Anglo-Saxon surnames project (Longley *et al.* 2007) suggested that there would be significant public interest in the classification at local scales. The 'e-Society profiler' part of the <http://www.spatial-literacy.org> website was launched on 1 August 2006 and immediately attracted some public interest alongside the surname profiling utility on the same site. A press release was duly prepared and sent to some of the national media outlets that had previously featured our surnames research. The release was picked up by BBC's online news service and was featured on its 'Technology' page on 8 August 2006 (Britain's digital tribes revealed: BBC News 2006 and Figure 2). The news feature included the following passage giving prominence to our desire to collect feedback on the classification:

SPATIAL-LITERACY.ORG
a centre for excellence in teaching and learning

Home Contacts People Projects Resources

Ever wondered how technologically enabled your neighbours are?

By **entering a postcode** in the e-Society profiler search box and **clicking "GO"**, our classification will tell you about the type of technology use in your neighbourhood and those areas with similar profiles within Great Britain.

This project has been developed by the Spatial Literacy team at UCL as an outcome of an earlier ESRC funded e-Society project.



- About the e-society
- What are the groups?
- Feedback

Enter a Postcode...

e.g. NW6 1CN **GO**

Before clicking "go" please read our data and data collection policy [here]

The E-Society profiler tool is solely for private, academic and public policy use only. Commercial use of these data is strictly forbidden. The E-Society profiler is not currently available under any commercial licence.

Figure 1. The user interface of the e-Society profiler.

Britain's digital tribes revealed

By Jonathan Fildes

Science and technology reporter, BBC News

Households in Britain can be classified into 23 "e-types" depending on their access to technology, say researchers.

E-types include mobile explorers, the e-committed and rational utilitarians.

The researchers, from University College London (UCL), say the profiles could be used to inform future policies on access to digital technology.

Every postcode in Britain has been assigned a classification which people can check online to see if they agree with the researcher's analysis.

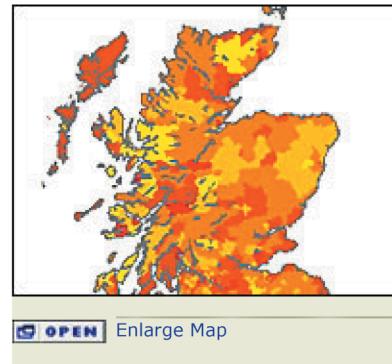
"What really emerges is that almost all of the types have some interaction with technology," said Professor Paul Longley, who led the study at UCL. "In a sense we are all digital now"

Figure 2. An extract of the national news feature arising from the consultation exercise (©BBC: news.bbc.co.uk/1/hi/technology/5256552.stm).

The team are now encouraging people to check their postcode and send the researchers feedback on their analysis. People who disagree with their assigned e-type can send ... their own classification from the list of eight groups and 23 types.

BBC News (2006)

Internet-based response solicitations have a number of disadvantages (e.g. Jones 1999), which are generally well known if not always successfully accommodated in PPGIS research: they exclude those who do not have access to the Internet; it is difficult to ascertain the identity of respondents with any confidence; they are vulnerable to multiple responses by a single individual; and they are vulnerable to wilful misrepresentations by self-selecting respondents. Yet, in our context, the BBC Online feature was a very useful medium through which to solicit feedback about our e-Society classification, because readership of an online technology feature by the UK's major broadcaster is very likely to garner the interest of individuals likely to be engaged with ICT-related activities. Thus analysis of the profile of responses was itself likely to provide a measure of the success of the classification. Moreover, BBC Online offers national coverage and has a readership likely to have a less overall political or other bias than that, for example, individual newspapers. It also attracts a readership that, whilst



engaged in ICT use, is border than those interested in new developments in technology for their own sake – as might be the case with an IT magazine feature, for example. Together, these criteria argue that the feedback is likely to be broadly representative of the socio-spatial patterns of engagement in ICT-related activities, albeit with a possible bias towards those interested in current affairs. However, the nature of the search and feedback facilities on the website did not enable us to document or restrict multiple searches, or indeed to record the number of searches that individual users actually made. It is thus perhaps most accurate to describe the patterns of searches as of ‘unit postcodes that are of interest to a broad spectrum of members of the e-Society’.

As an exercise in public consultation, the volume of responses suggests that we were very successful. The website attracted a total of 79,051 hits over the 13-day period that forms the basis to the analysis in this paper, with 20,694 hits on 8 August (the afternoon on which the feature was posted) and 22,113 hits on the following day. We believe that the magnitude of this response dwarfs almost all academic consultation exercises. The distribution of hits over the study period is shown in Figure 3. Figure 4 presents the number of postcode searches expressed as a percentage of the total number of unit postcodes per CAS ward. This map was created using the postcode searches that were undertaken using our site and, we suggest, presents a plausible representation of the geography of Great Britain’s ‘e-Society’. However, for reasons outlined above, the spatial concentrations of usage that may be identified using the map, such as those around Cambridge, may suggest not only heightened levels of ICT usage but also local patterns of high repeat usage of the facility.

However, a high absolute level of response does not, of course, mean that the consultation results are unbiased. Moser and Kalton (1985) differentiate between

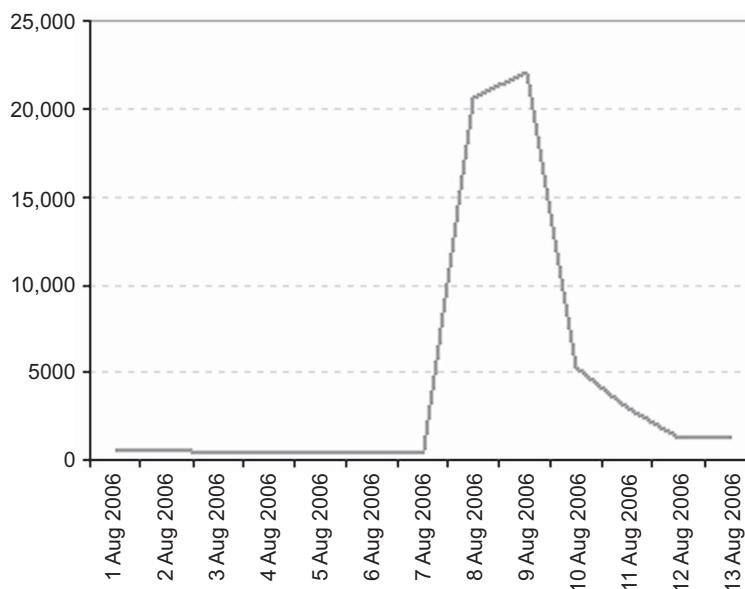


Figure 3. The frequency of unique visits to the site over the study period.

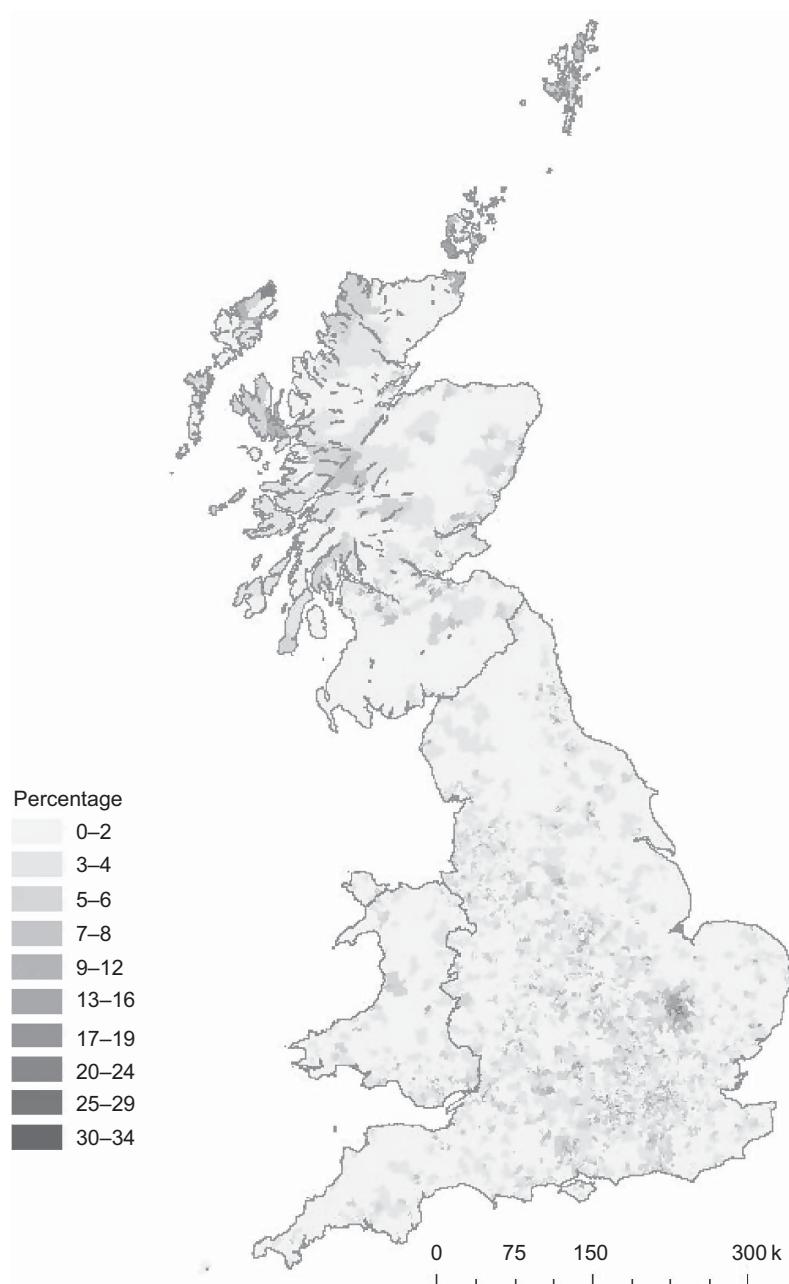


Figure 4. The percentages of unit postcodes within each CAS Ward that were searched during the study period.

internal validation of the results of social survey investigations, predicated upon assessment of the consistency and plausibility of the information collected, and **external** validation in which the quality and consistency of the information collected is assessed with reference to other data sources.

3.1 Internal validation

Our first internal validation measure arises from the observation that the medium of our consultation exercise, the national BBC Online news service, is very appropriate in improving our understanding of the geography of the ‘e-Society’. This being the case, we expect *a priori* that more searches will be undertaken of unit postcodes that we classify into ‘e-engaged’ Groups and Types.² Figure 5(a) shows the numbers of successful searches carried out on postcodes belonging to each Type of our classification (Table 1). In Figure 5(b), these data have been modified by subtracting from each Type the share of the searches that the Type would be expected to have given the percentage of GB postcodes assigned to it. This makes it possible to identify the Types which are over- or underrepresented in the searches – that is, a value of zero indicates that a Type was the subject of the same number of searches as would be expected. Scores above the zero line give the numbers of searches additional to expectations, which are balanced overall by the lower-than-expected assignments for other Types below the zero line³.

The observed composition of the searches conducted produces some validation of the success of the classification. In interpreting the results, it is important to remember that targeting applications of geodemographics present at best a very inexact science and that a ‘successful’ private sector application of a general-purpose classification might seek to, say, double the response rate to a mail shot from 2.5 to 5% – still representing a ‘failure’ rate of 95%. Postcodes assigned to Groups A and B (the ‘e-unengaged’ and ‘e-marginalised’) are, as would be expected in an Internet-based search, heavily underrepresented; postcodes from Group C (‘becoming engaged’) are up to one-and-a-half times more likely to be searched than expected, given the numbers of unit postcodes assigned to this Type; the within group difference for Group D identifies a surplus of enquiries about the residential postcodes of ‘small time net shoppers’ (Type D12) postcodes but a deficit for the ‘e for entertainment’ category (Type D13), which sees the net primarily as an entertainment rather than an information medium. The other (e-engaged) Types are all characterised by a surplus of searches, except for ‘rational utilitarians’ (Type E14) whose focused use of ICTs might suggest that they are less likely to participate in online surveys of this nature. Postcodes in the most engaged Group (Group H: “e-experts”) are very likely to be searched given the number of assignments to this Group.

These results demonstrate that the general-purpose Internet-based medium through which the feature was disseminated disproportionately reached a spectrum of users that can be considered core to the ‘e-Society’, while the underrepresentation of Groups that were predicted not to be ‘e-engaged’ also provides a validation measure. Because the feature was made available online to a national audience, and notwithstanding the caveats of small numbers of people making multiple searches, Figure 4 presents a plausible representation of the Great Britain geography of the ‘e-Society’.

In most private sector geodemographic systems, classification is taken as an outcome that cannot be disputed by the individuals who are classified⁴ – and indeed the less sophisticated indicators that are used in the public sector (such as the Index of Multiple Deprivation) are similarly closed to challenge if, for example, residents were concerned at the stigma associated with the ascription of ‘deprived’ status to their neighbourhoods. Our contention is that critiques of geodemographics need to refocus upon the unassailable and inert nature of classification outcomes rather than the laudable quest to build usable generalisations about spatial distributions. This

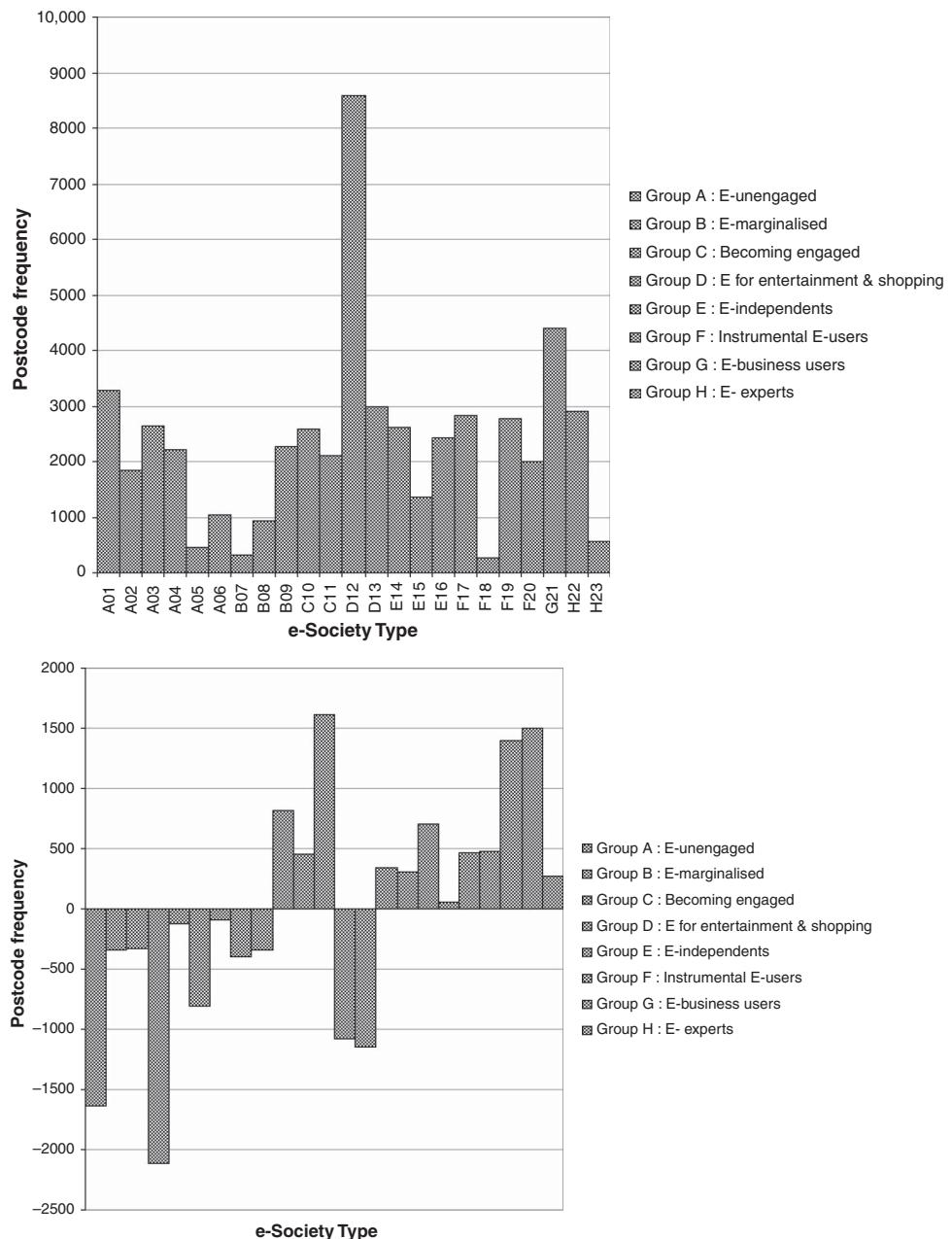


Figure 5. (a) The frequency of postcode searches during the study period; and (b) relative under- and overrepresentation of searches made by e-Society Type.

suggests a need to identify the sources of classification errors and to disentangle the ways in which they operate.

The approach that we adopted in this research was publicly to acknowledge the problems inherent in small area classification and to seek user feedback as a means of identifying, and thence correcting, possible systematic classification errors. 2707

Table 2. Comparison of e-Society predictions with user feedback evaluations.

	Feedback Group frequency & percentages								
	A	B	C	D	E	F	G	H	Total
A	28 (2.2)	47 (3.6)	179 (13.7)	215 (16.5)	294 (22.6)	158 (12.1)	59 (4.5)	322 (24.7)	1302 (100)
B	0 (0.0)	6 (2.9)	31 (14.8)	29 (13.9)	46 (22.0)	20 (9.6)	5 (2.4)	72 (34.4)	209 (100)
C	3 (1.5)	2 (1.0)	8 (4.0)	13 (6.5)	26 (12.9)	22 (10.9)	4 (2.0)	123 (61.2)	201 (100)
D	2 (0.4)	5 (1.1)	6 (1.3)	14 (3.1)	87 (19.3)	60 (13.3)	18 (4.0)	259 (57.4)	451 (100)
E	4 (2.2)	1 (0.5)	5 (2.7)	7 (3.8)	23 (12.4)	28 (15.1)	8 (4.3)	110 (59.1)	186 (100)
F	2 (1.2)	2 (1.2)	0 (0.0)	4 (2.4)	10 (6.0)	25 (15.1)	8 (4.8)	115 (69.3)	166 (100)
G	0 (0.0)	2 (1.3)	3 (1.9)	4 (2.6)	13 (8.4)	19 (12.3)	2 (1.3)	112 (72.3)	155 (100)
H	1 (2.7)	0 (0.0)	0 (0.0)	1 (2.7)	2 (5.4)	1 (2.7)	0 (0.0)	32 (86.5)	37 (100)
Unknown	34 (2.7)	31 (2.5)	46 (3.7)	116 (9.3)	203 (16.3)	172 (13.8)	75 (6.0)	568 (45.6)	1245 (100)
Total	74	96	278	403	704	505	179	1713	3952

(69%) of the 3952 feedback responses were usable, with the remainder largely pertaining to industrial or commercial premises. Table 2 compares the e-Society predictions with user feedback evaluations, aggregated for presentation purposes to the Group level⁵. Although sequencing of our Groups from A through H does not imply a 'ladder' of progressive engagement with ICTs, successive Groups do nevertheless exhibit greater awareness and use of them. With few exceptions (specifically the feedback that more postcodes be reassigned from Group G to Group E than vice versa), the scores and percentages in all of the elements above the principal diagonal of Table 2 are larger than those beneath it. This reveals the general outcome of the consultation to be the suggestion that unit postcodes be assigned to the more engaged Groups. Group H, the most ICT-literate Group, is the suggested destination for the greatest numbers of reclassified postcodes, which is perhaps the immediate reclassification for users that considered their selected postcodes to be engaged, but who could not be troubled to read the lengthy classification in its entirety.

The most obvious apparent misclassification of the e-Society appears to be the overly high assignment of individuals to Group A. 48% of all usable feedback pertained to postcodes that had been classified into this 'e-unengaged' Group. This might in part be expected, given that our classification was finalised in 2004 and that there have been shifts in the subsequent period in the overall level of engagement with ICTs. Of more concern, however, is that the suggested reassessments from Group A (and indeed those from the other less engaged Groups) to the 'higher' e-Society echelons are of broadly similar magnitudes. This suggests that misallocations are rather evenly spread across Groups B, C, D, E, F, and G. There may be an important issue for geodemographics research here, which underlines their blunt-edged power to discriminate between different groups throughout society (and supporting the arguments of Curry 1995, rather than Goss 1995, cited in our introduction).

A related issue that may be investigated through this feedback concerns the merits of updating of geodemographics. Some commentators (e.g. Sleight 2001) have observed that, for the vast majority of established neighbourhoods, temporal updating is not particularly important, since although the actual residents of neighbourhoods may well change over time, their dominant characteristics will not to any significant degree. This is clearly not the case in an absolute sense with regard to the e-Society, which is characterised overall by cumulative engagement as new technologies gain more widespread use. The rather even distribution of reported misclassifications across other categories appears to sustain our contention that each class defines a spectrum of usage rather than a rung on a predictable 'ladder' of engagement. It remains for future research to gauge the extent to which the spectrum of ICT usage within and between categories changes over time.

Figure 6 presents three facets to the feedback exercise. First, Figure 6(a) shows the amount of feedback received about postcodes falling into each e-Society Type (the 'origin' Type) – although this does not correct for the variable number of unit postcodes assigned to each Type of the classification. The composition of the feedback may be assessed by subtracting a score from each of these figures that represents the level of feedback that would be expected if feedback were directly proportional to the unit postcode-level distribution of Types. Using this calculation Figure 6(b) shows that there is a heightened propensity for users to leave feedback about postcodes classified as Group A (e-Unengaged).

In Figure 6(c) the emphasis is upon the distribution of assignments suggested in the feedback (the 'destination' Type), again standardised for the distribution of Types at the unit postcode level. The feedback clearly suggests the need to assign more postcodes to Types with higher levels of engagement with information technologies. Similarly, Table 2 also shows that feedback postcodes were overwhelmingly in high-tech categories – 78% in Groups E, F, G, and H, which account for a more modest 33% of the unit postcode assignments. Suggested assignments to the e-unengaged (Group A) accounted for less than 2% of assignments (but see below for discussion of data quality issues).

3.2 *External validation*

Moser and Kalton's (1985) second set of validation measures entail reference to external data sources in order to cross check results. Reference has been made in the preceding sections to commercial geodemographic systems, of which our e-Society classification can be thought of as a niche variant (albeit designed for public sector applications). The most popular commercial system in the UK, in terms of market share, is Experian's (Nottingham, UK) 'Mosaic' product, the 2002 version of which is based upon a mixture of 2001 Census data, public sector sources, and consumer surveys. It has a similar, if more elaborate, structure to the e-Society classification, comprising 61 Types that may be aggregated into 11 Groups (Experian 2007). Mosaic is marketed as a general-purpose geodemographic discriminator and claims no particular provenance with regard to awareness and use of ICTs. However, one of the often remarked characteristics of different geodemographic systems is that they tend to converge at the more aggregate (group) level where commonalities of social structure are likely to be more apparent. This being the case, Mosaic does provide a way of evaluating the e-Society classification in relation to a comparator with similar structure, which was devised using a similar classification methodology.

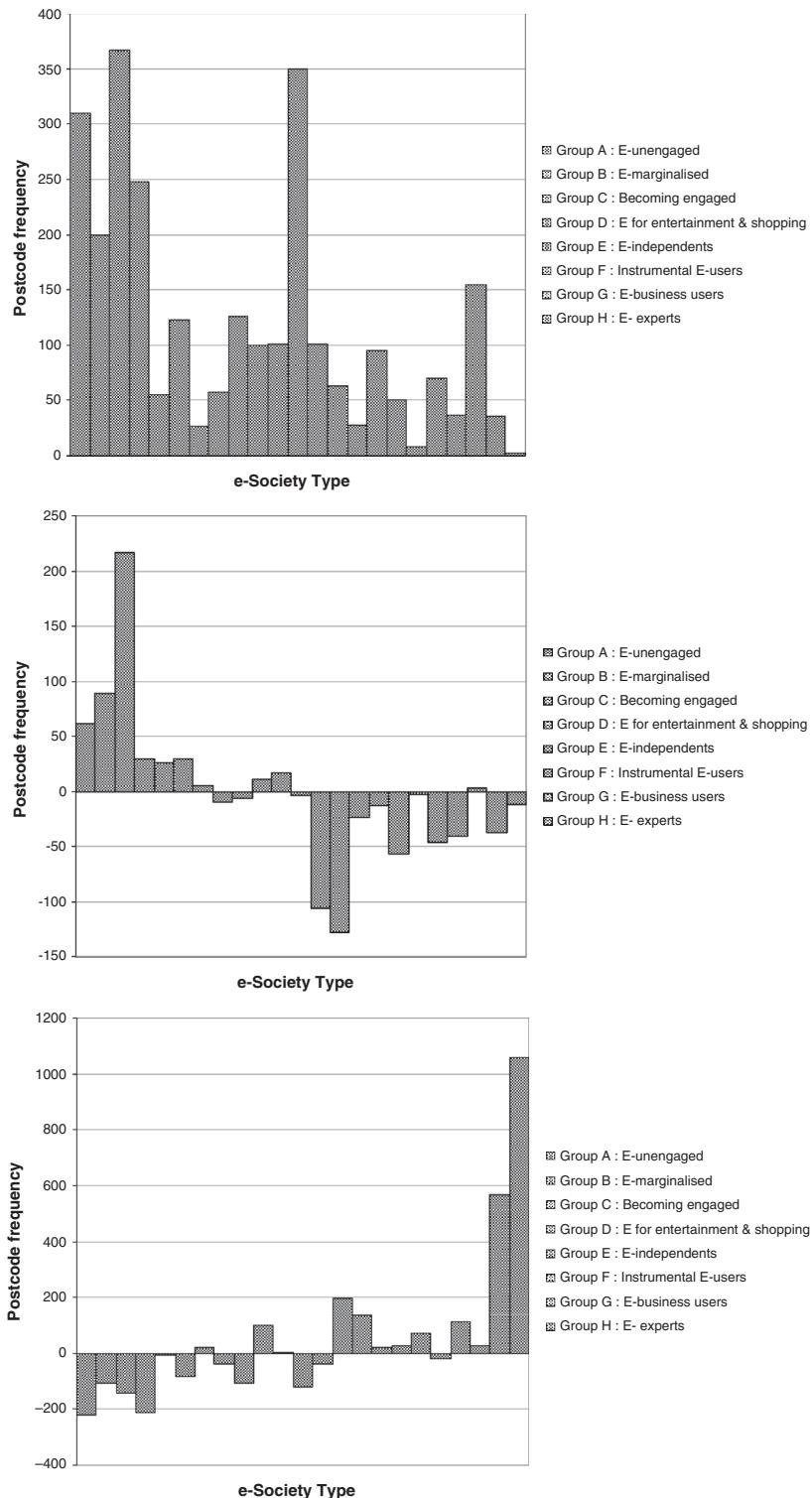


Figure 6. (a) Frequency of feedback according to origin e-Society Types; (b) frequency of feedback by origin e-Society Type; and (c) frequency of feedback by destination e-Society Type.

In order to facilitate this comparison, we adopt the geodemographics industry practice (e.g. Webber 2004a, 2004b) of creating index values. For any Mosaic Group, it is possible to create a series of index values for each e-Society Group using the equation:

$$s_{ij} = \frac{\sum_{i=1}^{t_{ij}/i=23} t_{ij}}{\sum_{i=1}^{b_{ij}/i=23} b_{ij}} \times 100 \quad (1)$$

where s_{ij} denotes the index score for e-Society Group i and Mosaic Group j ; t_{ij} was the number of searches undertaken on the e-Society Group i postcodes that are also classed within Mosaic Group j ; and b_{ij} is the total number of postcodes in e-Society Group i and Mosaic Group j in the classification. To this end, Table 3 presents index values for each of the 11 Mosaic Groups, where an index score of 100 identifies the number of searches on e-Society Group i to be proportional to its frequency in the Group as a whole, an index value of 200 identifies it as twice as common, an index value of 50 identifies it as half as common, and so forth. The results suggest the expected general convergence of the two different geodemographic classifications at the Group level: the 'E-unengaged' are disproportionately concentrated in low-status Mosaic Groups such as 'Twilight Subsistence' but also the aging 'Grey Perspectives'; the 'E-marginalised' are also socially marginal in the 'Welfare Borderline' Group; those 'Becoming Engaged' in e-Society terms are concentrated in Mosaic's 'Urban Intelligence', 'Welfare Borderline', and 'Ties of Community' Groups; 'E for Entertainment and Shopping' is a predominantly a lower middle income preoccupation in Mosaic terms; 'E-independents', 'Instrumental E-Users', and 'E-Business Users' are all concentrated in the most affluent social echelons. There is evidence for 'E-experts' in the educated 'Urban Intelligence' Group and also the more deprived (but sometimes geographically contiguous) 'Welfare Borderline' Group. There is, of course, a measure of circularity in this reasoning, in that similar techniques and types of background demographic data were used in each of these classifications: however, the level of correspondence between the two classifications does provide a measure of external validation, especially given that they were each created using different data.

Against this picture of general consistency, it is possible to investigate further patterning at the Type level in both classifications, as shown in Table 4 (and by

Table 3. Cross classification of e-Society and MosaicTM Groups.

	e-Society Group								
	A	B	C	D	E	F	G	H	
Mosaic Group	A Symbols of Success	74	18	21	42	210	226	114	63
	B Happy Families	10	35	68	151	29	192	267	49
	C Suburban Comfort	153	23	15	120	161	71	99	9
	D Ties of Community	85	216	136	191	28	27	50	43
	E Urban Intelligence	16	136	326	49	13	73	106	417
	F Welfare Borderline	87	451	260	40	4	23	19	224
	G Municipal Dependency	210	371	28	114	3	13	9	6
	H Blue Collar Communities	198	122	24	189	23	15	17	3
	I Twilight Subsistence	315	76	51	64	35	18	20	11
	J Grey Perspectives	270	63	61	49	81	42	46	37
	K Rural Isolation	91	24	10	21	419	124	52	10

Table 4. Cross classification of e-Society and Mosaic™ Types.

	A01	A02	A03	A04	A05	A06	B07	B08	B09	C10	C11
A01 Global connections	24	5	26	40	246	104	182	75	52	303	51
A02 Cultural leadership	31	18	85	40	328	48	342	18	12	43	29
A03 Corporate chieftains	5	0	57	11	79	0	91	0	0	4	0
A04 Golden empty nesters	16	19	331	12	183	8	78	5	0	0	0
A05 Provincial privilege	91	53	480	57	349	17	75	0	3	5	2
A06 High technologists	9	26	53	10	34	0	28	0	4	1	0
A07 Semi-rural seclusion	54	24	185	56	133	27	123	33	3	34	12
B08 Just moving in	14	12	0	10	0	44	0	0	10	44	32
B09 Fledgling nurseries	3	0	0	0	0	5	0	124	5	182	25
B10 Upscale new owners	4	2	2	4	29	4	13	33	2	17	4
B11 Families making good	12	13	16	5	0	3	64	49	6	54	46
B12 Middle rung families	16	86	5	12	8	6	52	25	6	14	13
B13 Burdened optimists	0	0	2	5	0	0	17	290	77	210	227
B14 In Military quarters	0	0	29	0	0	72	0	81	0	87	0
C15 Close to retirement	40	183	196	12	36	5	65	0	0	3	0
C16 Conservative values	377	602	405	96	231	16	25	0	0	0	2
C17 Small time business	169	142	191	131	115	59	100	35	11	20	19
C18 Sprawling subtopia	258	364	209	136	96	9	30	4	3	1	6
C19 Original suburbs	78	118	145	26	129	26	198	30	5	20	29
C20 Asian enterprise	0	38	5	13	31	13	85	60	304	64	156
D21 Respectable rows	80	68	14	77	55	81	102	121	189	74	197
D22 Affluent blue collar	184	524	36	147	29	50	80	14	6	8	6
D23 Industrial grit	138	127	11	145	8	63	45	98	152	14	63
D24 Coronation street	43	31	2	62	0	202	32	233	560	106	347
D25 Town centre refuge	47	37	22	78	22	607	381	145	169	286	393
D26 South Asian industry	0	0	0	11	0	23	0	104	1251	75	283
D27 Settled minorities	37	24	5	55	29	73	97	473	380	228	328
E28 Counter cultural mix	17	3	2	34	53	171	244	274	240	535	208
E29 City adventurers	8	2	1	10	0	48	9	50	51	395	150
E30 New urban colonists	8	5	17	10	74	47	121	149	107	294	217

(Continued)

Table 4. (Continued.)

	A01	A02	A03	A04	A05	A06	B07	B08	B09	C10	C11
E31 Caring professionals	19	8	6	15	25	32	102	76	270	223	529
E32 Dinky developments	9	0	2	5	0	35	0	331	119	472	354
E33 Town gown transition	24	3	6	28	35	54	63	72	106	319	356
E34 University challenge	28	20	27	25	0	17	55	39	56	224	213
F35 Bedsit beneficiaries	9	5	4	34	0	107	85	141	184	357	412
F36 Metro multiculture	42	27	0	107	45	350	362	746	682	199	251
F37 Upper floor families	12	14	0	107	0	328	161	709	688	183	340
F38 Tower block living	63	37	0	217	0	1841	420	296	302	106	387
F39 Dignified dependency	51	57	0	652	47	1302	384	338	294	81	197
F40 Sharing a staircase	0	189	0	94	0	733	850	450	672	27	196
G41 Families on benefits	0	13	0	22	0	183	73	1314	704	120	101
G42 Low horizons	120	246	4	315	0	490	142	163	500	0	16
G43 Ex-industrial legacy	321	183	4	785	0	494	129	159	185	8	10
H44 Rustbelt resilience	228	633	3	259	0	157	200	44	61	6	15
H45 Older right to buy	584	462	32	555	136	123	26	19	15	3	20
H46 White van culture	253	178	27	234	86	142	91	82	127	28	38
H47 New town materialism	32	80	0	55	0	149	26	270	357	23	49
I48 Old people in flats	122	36	0	1150	151	1603	0	144	29	26	31
I49 Low income elderly	324	219	27	504	163	336	158	89	91	32	107
I50 Cared for pensioners	109	24	17	1775	101	727	0	0	0	0	0
J51 Sepia memories	238	55	115	904	689	877	104	73	15	66	16
J52 Childfree serenity	184	45	74	203	560	579	189	178	86	178	129
J53 High spending elders	331	55	622	200	531	90	103	36	0	8	0
J54 Bungalow retirement	661	157	321	446	720	102	0	33	0	12	0
J55 Small town seniors	408	120	113	354	100	211	180	116	28	36	23
J56 Tourist attendants	337	0	193	403	0	204	260	46	19	131	0
K57 Summer playgrounds	84	38	392	125	312	33	0	37	0	0	0
K58 Greenbelt guardians	30	0	168	31	142	5	128	23	0	12	0
K59 Parochial villagers	114	0	300	89	74	31	100	53	7	16	19
K60 Pastoral symphony	74	5	196	40	242	34	191	48	0	17	4
K61 Upland hill farmers	48	0	264	43	286	61	97	102	0	0	0

A01 Global connections	18	13	42	117	53	66	43	514	921
A02 Cultural leadership	46	11	94	230	103	379	127	382	359
A03 Corporate chieftains	8	3	218	311	134	640	113	74	578
A04 Golden empty nesters	8	0	356	345	377	447	64	46	114
A05 Provincial privilege	43	5	175	224	343	302	40	63	63
A06 High technologists	117	12	73	437	173	436	57	100	16
A07 Semi-rural seclusion	74	19	224	146	160	217	169	244	27
B08 Just moving in	8	30	17	0	0	16	0	151	23
B09 Fledgling nurseries	45	21	2	8	0	6	20	367	39
B10 Upscale new owners	53	22	8	16	5	76	766	370	407
B11 Families making good	213	36	12	53	9	77	52	153	407
B12 Middle rang families	357	189	17	156	139	50	13	306	39
B13 Burdened optimists	184	82	6	4	0	0	0	325	39
B14 In military quarters	0	0	0	0	0	0	408	1655	39
C15 Close to retirement	128	29	73	459	474	175	50	31	39
C16 Conservative values	28	7	28	72	409	22	15	7	39
C17 Small time business	146	58	209	77	95	75	14	123	39
C18 Sprawling subtopia	163	53	15	106	191	69	37	17	39
C19 Original suburbs	205	29	49	150	113	202	122	75	39
C20 Asian enterprise	255	189	26	81	11	63	104	35	39
D21 Respectable rows	250	114	7	31	19	19	183	79	39
D22 Affluent blue collar	176	242	50	99	176	14	0	23	39
D23 Industrial grit	204	509	14	13	9	3	28	38	39
D24 Coronation street	122	282	0	4	2	0	20	38	39
D25 Town centre refuge	58	61	22	14	20	3	109	62	39
D26 South Asian industry	31	248	0	0	0	0	91	9	39
D27 Settled minorities	127	125	5	14	3	20	240	25	39
E28 Counter cultural mix	29	45	14	35	19	7	391	15	39
E29 City adventurers	6	9	8	9	6	2	445	22	39
E30 New urban colonists	79	25	49	36	11	52	261	108	39
E31 Caring professionals	125	52	7	8	8	5	84	57	39

(Continued)

Table 4. (Continued.)

	A01	A02	A03	A04	A05	A06	B07	B08	B09	C10	C11
E32 Dinky developments	54	16	2	4	0	6	91	128	39	120	392
E33 Town gown transition	52	24	2	8	4	4	175	33	77	76	585
E34 University challenge	13	12	14	13	22	32	137	39	18	78	815
F35 Bedsit beneficiaries	13	25	0	21	0	3	176	27	70	34	642
F36 Metro multiculture	21	112	4	0	8	3	224	7	39	22	94
F37 Upper floor families	23	219	5	0	5	5	0	52	0	9	54
F38 Tower block living	16	23	0	0	0	0	0	25	0	0	70
F39 Dignified dependency	19	35	0	0	9	0	0	23	10	5	43
F40 Sharing a staircase	0	347	0	0	0	0	262	25	0	0	24
G41 Families on benefits	22	294	0	0	0	0	138	24	33	24	0
G42 Low horizons	23	552	0	0	9	0	0	8	0	3	0
G43 Ex-industrial legacy	16	273	8	0	4	0	40	4	0	5	7
H44 Rustbelt resilience	40	654	9	0	13	3	0	3	0	4	0
H45 Older right to buy	21	168	26	19	39	0	32	15	4	10	0
H46 White van culture	150	361	18	19	57	18	32	29	11	23	7
H47 New town materialism	97	786	10	13	7	0	0	65	9	25	6
I48 Old people in flats	16	0	0	0	0	0	0	0	33	15	0
I49 Low income elderly	62	151	39	30	59	15	0	45	10	26	18
I50 Cared for pensioners	0	30	17	0	18	0	0	0	0	0	0
J51 Sepia memories	4	11	13	0	111	0	0	0	34	38	0
J52 Childfree serenity	54	36	33	34	51	40	156	61	93	69	113
J53 High spending elders	23	7	164	44	217	52	101	7	40	27	9
J54 Bungalow retirement	16	15	35	0	25	16	0	49	8	10	0
J55 Small town seniors	89	81	99	47	70	22	37	64	22	51	15
J56 Tourist attendants	59	28	243	156	52	30	0	61	21	39	29
K57 Summer playgrounds	12	0	948	102	42	24	391	100	52	16	0
K58 Greenbelt guardians	22	10	897	154	77	182	118	91	248	56	18
K59 Parochial villagers	34	27	821	66	53	86	0	135	37	67	3
K60 Pastoral symphony	21	15	985	66	69	113	34	87	169	47	0
K61 Upland hill farmers	11	11	1026	140	52	89	0	114	111	29	0

57

substituting Types for Groups in Equation (1)). Thus, for example, unit postcodes classified by Mosaic as being part of the 'Global Connections' Type are about four times less likely than average to be assigned to the A01 Type ('Low Technologists': index value 24) but about three times **more** likely to be assigned to the C10 Type ('e-bookers and communicators': index value 303). This disaggregation shows high levels of consistency at the Type level, but some interesting differences: the general correspondence between e-Society Group A (the 'e-unengaged') and Mosaic's 'Suburban Comfort' does not, for example, hold for the Mosaic 'Asian Enterprise' Type or the e-Society 'Elderly Marginalised' Type. The low correspondence at the Group level between the e-Society Group A ('the e-unengaged') and Mosaic's 'Welfare Borderline' Group conceals very strong representation of the e-Society 'Elderly Marginalised' Type across all of the Types that make up this Mosaic Group. These are but two examples of ways in which examination of Table 4 suggests new overarching hypotheses to account for the apparent patterns in this cross classification. In particular, it seems clear that demography confounds the apparent correlation between material wealth and e-engagement, with some elderly groups showing low rates of e-engagement even or particularly when wealth is taken into account. There is a sense in this comparison that the more conventional social hierarchy of the general-purpose Mosaic system is clustered around general indicators of affluence, while the e-Society classification successfully captures motivations to e-engagement that lie beyond the financial. The e-Society classification may thus be of use in decoupling motivation from ability to pay in studies of the evolving e-Society.

Our final comparison of the e-Society classification with Mosaic is to seek to verify the feedback received, given the reasons outlined above to presuppose that PPGIS data collected through an anonymous Internet site may be unreliable. Table 5 presents similar index scores to Table 4, but created using the feedback results according to Mosaic Type. (Note the high incidence of structural zeroes in cells where no reassessments were suggested.) Our interpretation of the index values contained in Table 5 suggests, or is guided by, the following hypotheses:

- (a) If we have misclassified a postcode that is actually unengaged as highly engaged, we are unlikely to get respondents to correct us – because they are unlikely to use the system. Thus feedback assignments of unengaged to engaged are likely to be well motivated.
- (b) Attempts to classify highly engaged postcodes as unengaged are more likely, *a priori*, to be mischievous (as, e.g., with the suggested reassessments to e-Society Type A01 'Low Technologists' of postcodes that fall within the 'Global Connections', 'Golden Empty Nesters' and 'Semi-Rural Seclusion' Types).
- (c) However, distrust in category (b) feedback may be reduced if there are trends within an isolated e-Society or Mosaic Type – as is apparent across e-Society Type A05 ('Too old to be bothered') or the Mosaic 'Town Gown Transition' Type.

These are all examples of the ways in which use of and feedback from our website may be used not only to suggest ways in which our classification might be improved but also to flag potential problems with the quality of some of our feedback data.

Table 5. Feedback analysis according to Mosaic™ Type.

	A01	A02	A03	A04	A05	A06	B07	B08	B09	C10	C11
A01 Global connections	196	0	0	0	0	686	125	0	0	58	123
A02 Cultural leadership	0	0	0	0	0	0	98	0	0	0	48
A03 Corporate chieftains	0	0	0	0	304	0	0	0	0	71	0
A04 Golden empty nesters	221	0	0	0	281	0	0	0	0	131	139
A05 Provincial privilege	0	0	0	0	272	0	0	0	120	32	67
A06 High technologists	0	0	0	0	0	0	117	0	0	27	58
A07 Semi-rural seclusion	209	0	585	0	133	0	0	0	0	62	33
B08 Just moving in	0	0	0	0	0	0	83	0	0	19	0
B09 Fledgling nurseries	0	0	0	0	0	0	0	0	0	0	0
B10 Upscale new owners	0	0	0	0	0	0	0	0	0	101	0
B11 Families making good	257	0	0	0	0	0	0	0	0	0	81
B12 Middle rung families	0	0	0	0	0	0	0	0	0	35	74
B13 Burdened optimists	0	0	0	0	0	0	0	610	0	131	0
B14 In Military quarters	0	0	0	0	0	0	0	0	0	0	0
C15 Close to retirement	0	0	0	0	0	0	92	149	0	64	181
C16 Conservative values	233	653	0	0	148	0	74	484	0	86	220
C17 Small time business	0	0	712	0	162	0	243	264	0	151	40
C18 Sprawling subtopia	0	1207	0	503	137	0	0	0	0	223	237
C19 Original suburbs	103	0	0	0	0	0	0	0	0	92	0
C20 Asian enterprise	0	0	0	0	0	0	0	1464	0	0	0
D21 Respectable rows	0	0	0	0	0	0	0	0	0	0	143
D22 Affluent blue collar	179	0	0	0	0	0	114	0	200	212	112
D23 Industrial grit	0	0	0	757	0	568	103	168	0	168	255
D24 Coronation street	0	0	0	0	0	0	0	0	930	246	392
D25 Town centre refuge	504	0	0	0	0	0	0	0	565	149	159
D26 South Asian industry	0	0	0	0	0	0	0	0	0	697	0
D27 Settled minorities	0	0	0	0	0	0	0	0	0	89	94
E28 Counter cultural mix	248	0	0	0	0	0	0	158	257	73	0
E29 City adventures	155	0	0	0	0	0	0	0	0	49	49
E30 New urban colonists	0	0	0	0	0	0	0	139	0	100	42

(Continued)

Table 5. (Continued.)

	A01	A02	A03	A04	A05	A06	B07	B08	B09	C10	C11
A01 Global connections	70	33	128	58	0	0	201	0	57	123	163
A02 Cultural leadership	110	0	145	159	166	95	210	179	67	48	94
A03 Corporate chieftains	57	40	104	106	0	0	163	279	140	150	84
A04 Golden empty nesters	53	0	144	163	0	0	0	193	32	276	116
A05 Provincial privilege	166	143	155	111	232	110	146	62	141	167	84
A06 High technologists	109	31	80	163	119	151	250	160	107	115	120
A07 Semi-rural seclusion	75	52	182	31	182	86	107	30	123	114	87
B08 Just moving in	140	131	85	39	57	108	179	114	153	102	155
B09 Fledgling nurseries	123	57	25	51	0	71	118	0	151	54	91
B10 Upscale new owners	27	38	83	34	49	187	78	133	67	71	120
B11 Families making good	61	86	93	190	0	159	88	75	188	80	90
B12 Middle rung families	56	39	51	105	0	48	80	0	103	147	165
B13 Burdened optimists	0	245	64	44	0	182	100	86	43	46	155
B14 In Military quarters	0	0	0	0	0	0	0	0	0	0	0
C15 Close to retirement	189	96	94	64	94	119	148	42	148	90	88
C16 Conservative values	250	156	118	138	203	144	159	102	68	18	56
C17 Small time business	76	64	157	57	110	131	87	185	56	139	89
C18 Sprawling suburbia	193	90	109	160	70	44	110	63	111	51	71
C19 Original suburbs	37	86	105	198	224	127	35	150	151	48	113
C20 Asian enterprise	336	0	205	418	0	0	0	0	0	0	62
D21 Respectable rows	90	0	88	135	33	156	104	177	133	95	106
D22 Affluent blue collar	128	89	91	185	116	184	61	156	157	56	102
D23 Industrial grit	39	189	94	120	35	167	111	0	95	51	114
D24 Coronation street	99	138	30	62	90	0	0	0	122	130	109
D25 Town centre refuge	180	84	219	75	219	104	0	0	148	0	88
D26 South Asian industry	374	0	0	0	0	0	0	0	0	0	69
D27 Settled minorities	107	50	22	222	65	185	0	88	176	47	105
E28 Counter cultural mix	30	124	108	37	107	51	0	0	218	155	130
E29 City adventurers	0	78	68	46	34	96	0	45	68	97	88
E30 New urban colonists	64	45	107	199	0	111	0	118	99	42	147
E31 Caring professionals	105	88	26	209	77	109	120	206	78	70	133

E32 Dinky developments	105	147	96	0	383	0	0	0	0	0	116	149
E33 Town Gown transition	105	196	43	0	0	0	0	0	0	129	46	168
E34 University challenge	0	131	57	116	0	0	0	0	0	115	0	172
F35 Bedsit beneficiaries	70	49	64	131	128	61	0	0	172	46	155	130
F36 Metro multicultural	160	0	49	100	146	0	0	0	0	105	0	140
F37 Upper floor families	0	588	0	0	383	0	0	0	259	552	0	46
F38 Tower block living	374	0	0	0	0	0	0	0	230	491	137	0
F39 Dignified dependency	210	0	128	261	191	0	0	0	129	0	39	0
F40 Sharing a staircase	0	0	171	0	0	0	686	0	0	0	103	61
G41 Families on benefits	0	118	51	105	306	145	0	0	0	0	62	92
G42 Low horizons	174	324	0	0	317	100	0	0	71	76	107	38
G43 Ex-industrial legacy	86	241	0	107	157	75	0	0	0	57	127	66
H44 Rustbelt resilience	129	241	53	54	0	0	0	0	106	113	63	38
H45 Older right to buy	70	294	64	44	0	303	100	172	129	0	103	84
H46 White van culture	92	207	68	138	101	128	159	45	114	73	68	105
H47 New town materialism	112	235	68	70	102	97	161	0	69	147	21	110
I48 Old people in flats	0	0	114	0	0	0	0	0	0	245	0	123
I49 Low income elderly	191	107	93	95	0	132	110	0	94	351	56	59
I50 Cared for pensioners	168	0	205	0	306	0	0	412	207	0	0	37
J51 Sepia memories	210	147	64	261	191	0	0	0	0	138	0	184
J52 Childfree serenity	113	136	98	101	118	84	93	119	60	85	131	99
J53 High spending elders	124	173	136	185	45	214	213	242	61	195	45	38
J54 Bungalow retirement	60	252	110	75	0	104	0	0	0	79	44	105
J55 Small town seniors	78	0	95	97	321	135	112	96	96	0	72	81
J56 Tourist attendants	0	294	256	0	191	0	301	0	129	0	193	46
K57 Summer playgrounds	140	196	256	0	255	242	0	0	0	0	103	61
K58 Greenbelt guardians	18	49	181	44	96	151	151	129	43	69	110	111
K59 Parochial villagers	92	0	186	114	0	53	175	0	75	201	79	94
K60 Pastoral symphony	33	0	201	41	0	0	189	161	0	130	158	87
K61 Upland hill farmers	0	102	178	0	0	126	210	0	192	0	27	112

4. Conclusion

The GIS literature remains confused about geodemographics. Pickles' (1995) influential collection of essays opened up this debate but in hindsight presents an ambiguous, even contradictory, position in which, on the one hand, geodemographics is deemed to provide inaccurate depictions of socio-economic conditions while, on the other, it is suggested that geodemographic classifications threaten to rewrite entire landscapes of consumption. The commercial imperatives of the geodemographics industry dictate that rather little has been done to resolve this apparent contradiction: commercial solutions are by their very nature not 'open source' and there is a lack of transparency, not least with regard to the weights that are assigned to the apparently arbitrary ingredients of geodemographic cocktails. However, even where the methodology and data used are clearly documented in the literature, the disjuncture between protagonists and detractors of the approach remains because the techniques are specialist, and the pen portraits are necessarily generalised and reductionist.

In this paper, we have sought to move the debate about the use of geodemographics on from confrontational but unsubstantiated statements about what might and what might not be possible using these systems, towards a measured empirical evaluation of how such systems appear to work in practice. In so doing we concur with Harvey (2006) that issues of trust in GIS are key to building support for spatial data infrastructures, particularly when the public are themselves stakeholders in the applications in which geodemographics are applied. We suggest that the consultation process developed here presents a valuable route beyond this impasse. In this conclusion, we begin by extending the technical arguments of Longley and Goodchild (2008) by reflecting on the necessary prerequisites to wider acceptance of geodemographics research in policy-relevant applications. First, it is clear that geodemographic classifications need to be much more transparent (and hence scientifically reproducible) with regard to method of classification and the provenance of the data used to create them. Second, we recommend that wherever possible, it is important that the results of classifications be made available in the public domain, in order that individuals and organisations may investigate how they and others that they represent have been classified. There is a need to develop new business models that are capable of reconciling best scientific practice with commercial imperatives.

To these major requirements, we add a suggestion, arising from the analysis presented here, that stakeholders in the classification – specifically those who are classified by it – should be provided with portals through which to comment on the outcome of the classification process. There are strong reasons for suggesting that modelling systems engineered at the level of the individual are inherently superior to those founded upon aggregations (Hensher and Johnson 1981), and the innovation of PPGIS arises out of recognition that it is at the individual level that consultation and hence validation should occur. This third comment is in the interests of those creating classifications, since it simultaneously offers the prospect of correcting individual anomalies (arising, for example, through brown field redevelopment or regeneration initiatives) and of revisiting more fundamental issues of classification (such as the inclusion and weighting of particular variables).

The major caveat to all of this is respondent incentivisation and validation of science through surveys that are themselves inherently unscientific. Specifically, social surveys too often assume ubiquity of online media and full engagement with these media of everyone with an interest in contributing to surveys. We have been fortunate

in the conduct of our own consultation, in that our communications channel and response medium were core to our investigation and thus were of direct relevance in targeting the segments of the population that were of greatest interest to the research. Our findings are nevertheless vulnerable to the other criticisms of social survey research (De Vaus 2002), namely that is unclear exactly who is responding, how often, and on behalf of whom.

The only technical response to such problems entails recourse to external validation. In our own example, comparison of our results with the Mosaic geodemographic classification provides some limited attempts to hijack the consultation process through feedback that seems improbable, even absurd. Our own experience of this, however, provides grounds for some optimism that geodemographics itself can provide appropriate checks on validating the veracity of feedback received. This is a pragmatic response, of course: we seek to validate our classification using a commercial solution that is of unknown provenance (but see Webber 2004b) and, detractors might argue, is founded upon similar misguided attempts to generalise about socio-economic distributions. Yet the general consistency thrown up by our external validation exercise does suggest that geodemographic classifications provide an important way of evaluating feedback and should be integral to studies of the 'e-Society' (Elwood 2002; Corey and Wilson 2006). Few generalisations about socio-economic distributions are founded upon incontestable facts: we believe that this study exemplifies the ways in which consultation can be used to reinforce and substantiate the inductive classification upon which geodemographics depends, while building trust (Harvey and Tulloch 2006) amongst the broadest constituency of citizens and stakeholders.

The broader context to these technical issues is that there is a need for academics and policy makers to understand the potential and limitations of geodemographic systems, both to guide informed debate and thence to guide the public. However, it is extremely unlikely that the outcome of informed debate will ever be an unequivocal and universal statement as to whether geodemographics makes a valid, ethical and useful contribution to the analysis of socio-economic systems: indeed, we suspect that the most persuasive arguments are likely to be quite domain or even application specific. Yet we believe that the public use of our own classification and the volume of feedback received has demonstrated widespread public interest in taxonomies of the public and that the nature of the feedback suggested that our particular classification was fit for the purpose for which it was designed. Our consultation exercise suggests that there is much to be gained from greater public outreach and much public support available to build better geodemographic classifications.

In closing, we suggest that this paper has opened up some routes towards a more enlightened approach to geodemographics. First, our research project has provided evidence that effective bespoke geodemographic classifiers can be devised with particular applications in mind. The wide availability of software suitable for creating geodemographic applications, coupled with the vastly improved access to devolved data holdings across the Web, potentially empowers a huge range of interest groups to create classifications that are honed to particular applications and contexts. It is clear that there is no longer any need for reliance upon a small number of proprietary, 'black box' geodemographic systems, and we see this as a very positive development. If, in particular, the best profiles of the public can lead to better resource allocation of public goods in areas such as health, education and policing (Longley 2005), then everyone will surely benefit.

Second, the same technologies that are making neighbourhood data available to ever wider constituencies of geodemographic system developers are also empowering individuals to access the results of such systems. The availability of CACI's ACORN system for individual address queries on the www.upmystreet.com website provides one example, and the site described in this paper provides another. As this movement gathers pace, so it is likely that GeoWeb 2.0 architectures will be increasingly used to make these results available to the wider public in ways that are readily intelligible to all but the most spatially illiterate. Third, and building upon this second point, the development of GIS-based web services offers the prospect of informing the wider public of the ways in which they are viewed by policy makers and marketers. This is something that those residents of neighbourhoods variously described as deprived or desirable should be empowered to find out; but it remains frustratingly difficult for even quite expert users of existing government statistics websites. Taken together, this more enlightened approach to geodemographics presents a valuable means of improving profiling of the public, for the public, and ultimately by the public.

Acknowledgements

This research was funded by ESRC Grant RES-335-25-0020 and supplementary funding from the ESRC e-Society Programme Director Roger Burrows. The website www.spatialliteracy.org is part of the HEFCE-funded 'Spatial Literacy in Teaching' (Splint) Centre of Excellence in Teaching and Learning. The authors are very grateful to Experian Ltd. for the supply of data to the project.

Notes

1. www.york.ac.uk/res/e-Society/
2. The BBC news feature did not make clear that the classification was only of residential postcodes and consequently some searches, possibly made by office workers during business hours, concerned non-residential postcodes. Other 'invalid' postcodes included those of recent inception (e.g. identifying new build properties) and those that had been changed in the period subsequent to creation of our classification.
3. A common practice in commercial geodemographics is to use index values instead of absolute frequencies in order to gauge the over- or underrepresentation of population characteristics. We do not adopt this practice here, however, because the consultative nature of our investigation requires that we do not accept that the total assignments of postcodes to Types are inevitably correct.
4. An exception in the UK is CACI Ltd.'s ACORN system, which is available to public access through the www.upmystreet.com neighbourhood profiling site, responses from which are used to refine the system (John Rae, personal communication).
5. Elements along the principal diagonal of the main table identify feedback responses where respondents have identified a different Type within the same modelled Group as being more appropriate. As such this 4.7% of usable responses can be considered to be 'near misses'.

References

BBC NEWS, 2006, Britain's digital tribes revealed. Available online at: news.bbc.co.uk/1/hi/technology/5256552.stm (accessed 22 January 2007).

BURROWS, R., and ELLISON, N., 2004, Sorting places out? towards a social politics of neighbourhood informatisation. *Information, Communication and Society*, 7, pp. 321-336.

COREY, K. and WILSON, M., 2006, Urban and Regional Technology Planning. *Planning Practice in the Global Knowledge Economy* (Ann Arbor: Michigan State University Press).

CURRY, M., 1995, On the inevitability of ethical inconsistency in geographic information systems. In J. Pickles (Eds), *Ground Truth*, pp. 68–87 (New York: Guilford Press).

DE VAUS, D.A., 2002, *Social Surveys*. (London: Sage).

ELWOOD, S., 2002, GIS use in community planning: a multidimensional analysis of empowerment. *Environment and Planning A*, **34**(5), pp. 905–922.

EXPERIAN, 2007, Mosaic UK. Available online at: www.businessstrategies.co.uk/upload/downloads/mosaic%20uk%20brochure.pdf (accessed 2 January 2007).

GOSS, J., 1995, Marketing the new marketing: the strategic discourse of geodemographic information systems. In J. PICKLES, *Ground Truth* (Eds), pp. 130–170 (New York: Guilford Press).

HARRIS R, SLEIGHT, P. and WEBBER R., 2005, *Geodemographics, GIS and Neighbourhood Targeting*. (Chichester: Wiley).

HARVEY, F., 2003, Developing geographic information infrastructures for local government: the role of trust. *Canadian Geographer*, **47**, pp. 28–36.

HARVEY, F. and TULLOCH, D.L., 2006, Local government data sharing: evaluating the foundations of spatial data infrastructures. *International Journal of Geographical Information Science*, **20**(7), pp. 743–768.

HENSHER, D.A. and JOHNSON, L.W., 1981, *Applied Discrete Choice Modelling* (London: Croom Helm).

JONES, S., 1999, *Doing Internet Research: Critical Issues and Methods for Examining the Net* (London: Sage).

LONGLEY, P. A., 2005, A renaissance of geodemographics for public service delivery. *Progress in Human Geography*, **29**, pp. 57–63.

LONGLEY, P.A. and GOODCHILD, M.F., 2008, The use of geodemographics to improve public service delivery. In J. Hartley, C. Skelcher, C. Donaldson, G. Boyne (Eds), *Managing Improvement in Public Service Delivery: Progress and Challenges*, pp. 176–194 (Cambridge: Cambridge University Press).

LONGLEY, P.A., WEBBER R., LI, C., 2008, The UK geography of the e-Society: a national classification. *Environment and Planning A*, **40**, pp. 360–382.

LONGLEY, P.A., WEBBER, R. and LLOYD, D., 2007, The quantitative analysis of family names: historic migration and the present day neighbourhood structure of Middlesbrough, United Kingdom. *Annals of the Association of American Geographers*, **96**, pp. 31–48.

MOSER, C.A. and KALTON, G., 1985, *Survey methods in social investigation*, 2nd edition (Aldershot: Gower Publishing Company Ltd).

PICKLES, J., 1995, *Ground Truth: the Social Implications of Geographic Information Systems* (New York: Guilford).

SLEIGHT, P., 2001, *Targeting Customers. How to use Geodemographic and Lifestyle Data in Your Business*, 2nd edition (Exeter: NTC).

WEBBER, R., 2004a, Designing geodemographic classifications to meet contemporary business needs. *Journal of Interactive Marketing* **5**(3), pp. 219–237.

WEBBER, R., 2004b, The relative power of geodemographics vis a vis person and household level demographic variables as discriminators of consumer behaviour. *CASA Working Paper*, 84 (London: UCL). Available online at: www.casa.ucl.ac.uk/publications/workingPaperDetail.asp?ID=84 (accessed 22 January 2007).