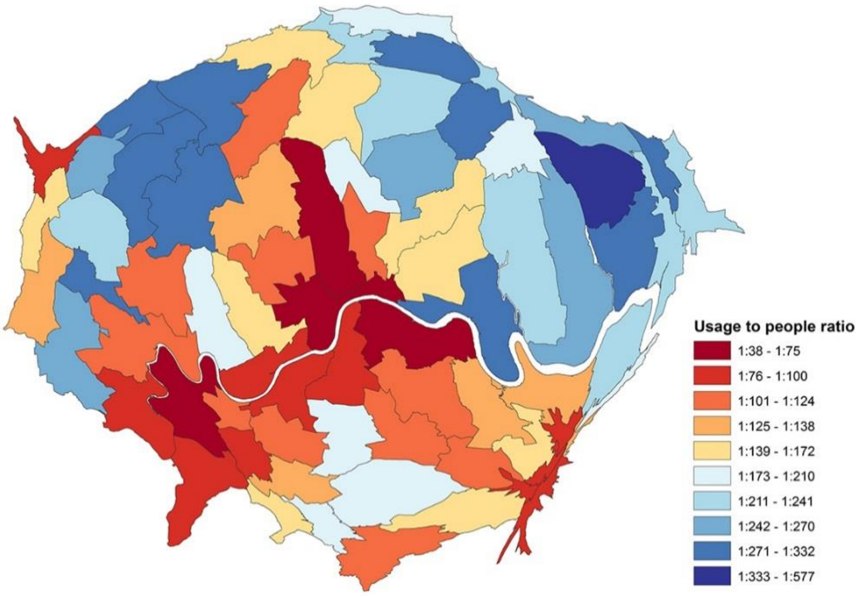
British Geological Survey

NATURAL ENVIRONMENT RESEARCH COUNCIL

A methodology for conducting a study into the communicative effectiveness between choropleth maps and cartograms.

Data & Science Services Programme

Internal Report IR/16/021



BRITISH GEOLOGICAL SURVEY

DATA & SCIENCE SERVICES PROGRAMME

INTERNAL REPORT IR/16/021

A methodology for conducting a study into the communicative effectiveness between choropleth maps and cartograms.

Holbrook, H W & Cartwright, C E

Keywords

Report; visualisation; cartograms; choropleth maps; results; effectiveness.

Front cover

Cartogram showing iGeology app usage in Greater London. Contains Ordnance Survey data © Crown copyright and database right 2014; National Statistics data © Crown copyright and database right 2014.

Bibliographical reference

HOLBROOK, H W & CARTWRIGHT, C E. 2016. A methodology for conducting a study into the communicative effectiveness between choropleth maps and cartograms.. *British Geological Survey Internal Report*, IR/16/021. 26pp.

Keyworth, Nottingham British Geological Survey 2016

BRITISH GEOLOGICAL SURVEY

The full range of our publications is available from BGS shops at Nottingham, Edinburgh, London and Cardiff (Welsh publications only) see contact details below or shop online at www.geologyshop.com

The London Information Office also maintains a reference collection of BGS publications, including maps, for consultation.

We publish an annual catalogue of our maps and other publications; this catalogue is available online or from any of the BGS shops.

The British Geological Survey carries out the geological survey of Great Britain and Northern Ireland (the latter as an agency service for the government of Northern Ireland), and of the surrounding continental shelf, as well as basic research projects. It also undertakes programmes of technical aid in geology in developing countries.

The British Geological Survey is a component body of the Natural Environment Research Council.

British Geological Survey offices

BGS Central Enquiries Desk

Tel 0115 936 3100 Fax 0115 936 3200
email enquiries@bgs.ac.uk

Environmental Science Centre, Keyworth, Nottingham NG12 5GG

Tel 0115 936 3100 Fax 0115 936 3200
email sales@bgs.ac.uk

The Lyell Centre, Research Avenue South, Edinburgh EH14 4AP

Tel 0131 667 1000 Fax 0131 668 2683
email scotsales@bgs.ac.uk

Natural History Museum, Cromwell Road, London SW7 5BD

Tel 020 7589 4090 Fax 020 7584 8270
Tel 020 7942 5344/45 email bgslondon@bgs.ac.uk

Columbus House, Greenmeadow Springs, Tongwynlais, Cardiff CF15 7NE

Tel 029 2052 1962 Fax 029 2052 1963

Maclean Building, Crowmarsh Gifford, Wallingford OX10 8BB

Tel 01491 838800 Fax 01491 692345

Geological Survey of Northern Ireland, Department of Enterprise, Trade & Investment, Dundonald House, Upper Newtownards Road, Ballymiscaw, Belfast, BT4 3SB

Tel 028 9038 8462 Fax 028 9038 8461

www.bgs.ac.uk/gsni/

Parent Body

Natural Environment Research Council, Polaris House, North Star Avenue, Swindon SN2 1EU

Tel 01793 411500 Fax 01793 411501
www.nerc.ac.uk

Website www.bgs.ac.uk
Shop online at www.geologyshop.com

Foreword

This report is the published product of a study by the British Geological Survey (BGS). The project involved establishing an overview of the communicative effectiveness of both cartograms and choropleth maps through an empirical test using various members of BGS staff as participants. Test examples were created and presented along with an evaluation as to their effectiveness and efficiency in communicating spatially orientated datasets. The authors are indebted to all the participants whom gave up their time and energy to the study, whom without their contribution, this report would never have materialised.

Acknowledgements

In addition to the BGS staff acknowledged in the Foreword, the authors are indebted to Wayne Shelley for his help and advice in procuring the BGS mobile app usage data. The authors would also like to thank Emma Bee and Kathryn Lee for editing this report.

Contents

FIGURES

Summary

In recent years, both cartograms and choropleth maps have grown in popularity as a chosen method of communicating spatial data, often in the context of an infographic. However, little is known about the effectiveness of these types of maps as a communicative tool. This report is an empirical study designed to firstly identify what method of visualisation is more effective in communicating spatial data between two types of map; and secondly, establish any rules associated with the type of data to be visualised, that those communicating spatial data must consider when choosing the type of map to use. Following the controlled test, significant findings were recorded, showing a clear difference in communicative effectiveness between the two map types. Regardless of the participant's personal preference, our research shows that the area cartogram is the most effective method of communicating a simple message, whereas choropleth maps are more successful at simultaneously communicating opposing values.

This report describes the methodology behind a trial that was designed to measure the communicative effectiveness between two common formats of data visualisation. The first part of the report introduces the background behind this area of research, followed by the methodology employed to help identify the strengths and any weaknesses in their communicative power. The results obtained from this research have propelled our understanding of communicative effectiveness beyond any work previously published; establishing guidelines on how and when such visualisations should be used, along with controversial findings showing the preferred choropleth map to be less effective than the cartogram.

1 Introduction

Choropleth maps and cartograms have witnessed a growth in popularity in recent years, due in part to the rise in Big Data (Cukier & Mayer-Schoenberger, 2013), improved access and availability of GIS (Goodchild, 2010; Tobler, 2004), and a growing trend in the use of infographics (Ferreira, 2014). Maps are frequently used as the vehicle of choice for the expression of spatially orientated messages as infographics (Guardian, 2014). Although choropleth maps and cartograms are a popular method of spatial visualisation, there is little published evidence demonstrating the communicative effectiveness of these maps (Nusrat & Kobourov, 2015). The criteria behind the use of such maps generally fall on either tradition (choropleth maps) or style (cartograms) when choosing a map type to communicate a desired message. Little, if any, emphasis is placed on the appropriateness of the map type being used when communicating spatial data (Griffin, 1983; Robinson & Petchenik, 1976). This report will explore and measure the level of communicative effectiveness between these two map types, with the primary aim of informing those communicating spatial data which method of visualisation is the most appropriate to use. Both types of map have the potential to display quantitative spatial data, and previous studies (especially for cartograms) have demonstrated this when they are presented alongside an inset map, and a clear annotated marginalia (Dent, 1975; Griffin, 1983; Kaspar, Fabrikant, & Freckmann, 2011; Nusrat, Alam, & Kobourov, 2015; Sun & Li, 2010). However, we have observed that the practice of using a magnitude key alongside an inset map in most published examples is generally overlooked. If magnitude is represented in any way beyond the weighted distortion of the polygon regions, it tends to be via the use of a mesh/grid (B. Hennig, 2015); a process that involves overlaying the original base map with a wireframe mesh prior to distortion (Guardian, 2014; B. D. Hennig, 2011; Weber Reuschel, Piatti, & Hurni, 2014). We can only assume that the main reason for not using an inset map, alongside a cartogram, is due to a desire by the author to keep the map as clear and concise as possible; reducing any potential for visual clutter.

Most measured tests between cartograms and choropleth maps prior to this report have used similar methods of portraying magnitude. Measurement of effectiveness between choropleth maps and cartograms have relied on areal size change to represent magnitude, whether they be proportional circles, bars, squares, or distorted abstract regions. Magnitude represented by areal ornament alone is less common. Based on this observation, we wish to measure the communicative effectiveness of choropleth maps and cartograms purely on the participant's ability to read the maps in their own right, without the added quantitative visual aids advised in previous papers. In other words, test the communicative effectiveness of these maps without the additional proportional symbols, magnitude keys, or inset maps previously deemed necessary for communication. The map types chosen for this test are the standard choropleth map (without added proportional symbols), and the 'Gastner and Newman' diffusion based area cartogram. The decision to move away from using proportional symbols for the choropleth map was necessary in order to ensure that the two types of map were truly distinct in their portrayal of magnitude. In order to achieve this, the choropleth map was designed to communicate categories of magnitude by the ornament shown for each area (in this case: colour), with the aid of a key; and the cartogram would show its magnitude via proportional changes in areal coverage size (shape distortion).

1.1 BACKGROUND

Published research on the communicative effectiveness of cartograms and choropleth maps, beyond algorithm variants (e.g. distortion methods of contiguous area cartograms) and visualisation methods (the style of area cartogram: e.g. Circular, Rectangular, contiguous, non-contiguous, etc.), is limited. Research by Dent (1975) measured the readability of magnitude (i.e. quantitative value) from various scaled types and style of region (i.e. polygon). This research found abstract symbols (e.g. square polygons) were most effective when displayed with a corresponding key. He went on to measure the effectiveness of communicating population between a generalised cartogram and a standard geographic/planimetric map. Dent's study used maps that were augmented by circles using 'absolute scaling' showing the level of population. Results revealed that the participants reading the scaled circles tended to overestimate magnitude, whereas the results from the non-contiguous cartogram showed an underestimation of magnitude. Subsequent research by Griffin (1983) measured the effective readability of both magnitude, and location, by adopting Dent's method of showing a standard planimetric map depicting electoral regions alongside a generalised contiguous electorate population cartogram. Emphasis was placed on measuring the speed and accuracy achieved by participants using what is now termed as 'linking', where readers, for example, simultaneously use a geographic map as a spatial referencing aid to an unlabelled cartogram, and vice versa (Nöllenburg, 2007); this method has since proved to be an effective way of communicating onscreen digital data (Buja, Cook, & Swayne, 1996; Dykes & Unwin, 1998). Griffin's results revealed that it took participants longer to identify standard unlabelled planimetric regions from labelled cartogram regions. Griffin also observed that participants adapted to the cartogram significantly by their third attempt (Griffin, 1983). By the mid-1990s, technical developments in both computing power and software opened the way for the development of new cartogram algorithm's, provoking a renewed interest in cartograms (Daniel Dorling, 1996; Gastner, Shalizi, & Newman, 2005; Wolf, 2005), along with theories on how these maps would be interpreted by readers (Openshaw & Alvanides, 2001; Speckmann, 2006; Tobler, 2004). By 2010, new systematic evaluations of the communicative effectiveness between cartograms and choropleth maps appeared. An online survey, created by Sun and Li (2010) enabled participants to compare various cartogram types alongside a choropleth map. On average, results revealed choropleth maps to be the preferred format by the participant, supporting Dent's findings (1975). Sun and Li went on to record which of the area cartograms used in the test proved most popular with participants. Results showed the 'Gastner and Newman' diffusion type (i.e. the Density Equalizing map) proved to be the most effective; leaving the Dorling circular type ranked as least effective by the participant (Sun & Li, 2010). Kaspar et al. (2011) measured accuracy and response times taken to read choropleth maps, augmented with proportional 'absolute scaling' circles, against 'Gastner and Newman' diffusion type cartograms, based on two distinct types of base map area: the regular grid like counties of Kansas, USA; and irregularly shaped regions of the Canton of Basel. Results showed that participants found cartograms, which had more regularly shaped regions, were quicker to identify, but not necessarily more effective to read. Cartograms with irregular shaped regions proved more effective at communicating area magnitude than the more regular (i.e. square) shaped regions. It was observed that processing regular shapes through the contiguous diffusion method distorted the shape too dramatically for the participant to make an accurate estimation of magnitude. It was further concluded that the contiguous diffusion method of the cartogram was better suited to irregular shaped regions. It was suggested that other methods of cartogram should be considered for the more regular shape type base map. The findings by Dent (1975) suggest that the non-contiguous method of distortion for regions of a more regular shape may prove more effective. Overall, there was insignificant difference in communicative effectiveness between the two styles of map, although, the more complex the question became, accuracy was higher for the choropleth map (Kaspar et al., 2011). Continuing with the online 'linking' method of map reading, Nusrat et al. (2015) focussed on measuring effectiveness in terms of time and accuracy, between four different types of area cartogram. Based on overall performance and qualitative results obtained from a questionnaire submitted by participants after the test, the Dorling circular cartogram and the contiguous area cartogram proved to be the most effective types. Interestingly, this result goes contrary to findings by Sun and Li (2010) whom placed the Dorling circular type cartogram as least effective. We believe this shift in the result is likely to be due to an increased scope of questions employed in the test, allowing for highlighting strengths and weaknesses for each type of cartogram to a greater degree. Close analysis of these 'task' based results demonstrate the Dorling circular type cartogram performed best with readers analysing and comparing trends; and the contiguous cartogram performed best in the task areas of change detection, area identification, and communicating min-max values. As for the remaining cartogram styles, the non-contiguous cartogram proved less effective, leaving the rectangular cartogram as the least effective, in both qualitative and quantitative measures of performance (Nusrat et al., 2015).

1.2 CHOROPLETH MAPS

Choropleth maps are a class of quantitative thematic map that traditionally use a sequence of shades or patterned fills to represent a group, or class-range, of values assigned to each defined regional area (Cuff & Mattson, 1982). Choropleth maps have been around since the early nineteenth century, but 1938 witnessed the earliest reference to the word 'choropleth', a term coined by the then Director of the American Geographical Society, John K. Wright (Crampton, 2011). Unlike the cartogram, the choropleth map retains its projection and topology. The apparent ease of reading choropleth maps have earned them widespread appeal (Dent, 1990) and are used to communicate all manner of quantitative information, for example, maps showing: election results, gross domestic product, world mineral production, etc.

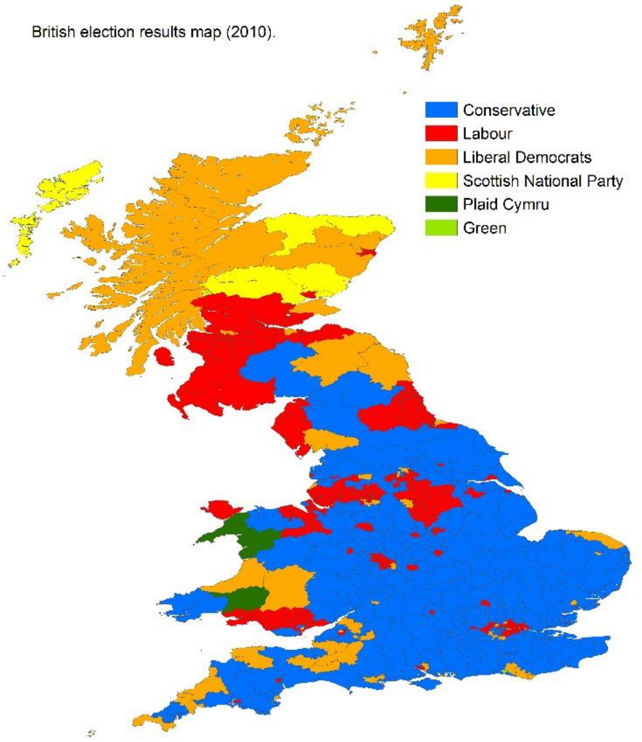


Figure 1: Choropleth map showing the 2010 UK Election results.

1.3 CARTOGRAMS

There are two types of cartogram: the distance cartogram, and the area cartogram (Shimizu & Inoue, 2009; Sun & Li, 2010). Distance cartograms show value over distance, which can be displayed in the form of points and vertices. Unlike the isochronic map used to show time over distance from a common starting point (Galton, 1881), the distance cartogram applies the value over distance rule to all inter-connected points simultaneously. This type of cartogram can be effective in communicating temporal movement information; for example, the overall time taken to travel between cities via a national rail network (Shimizu & Inoue, 2009).

An area cartogram is a map that distorts its regions according to a chosen value, displaying statistically aggregated data in a non-projected spatial context. Topology is secondary to the quantitative value assigned to each areal unit (region) and any projection associated with the original base map is completely discarded; as a result, some do not consider the cartogram as a map (Bortins & Demers, 2002). The balance between retaining original spatial shape (Keserica, Sučić, & Mihajlović, 2009; van Kreveld & Speckmann, 2004) and allowing a value to distort a region can vary depending on the subject matter (Kocmoud, 1997). Consideration between the level of shape retention and distortive freedoms are governed by the cartogram's ability to communicate its message.

An area cartogram can generally be classified into one of three sub-types: contiguous, non-contiguous, and the circular cartogram. Contiguous area cartograms use algorithms that allow each areal unit, or region, to be distorted by a value without breaking up the topology. Retaining the topological relationship can help the user identify regions, albeit sometimes heavily distorted. Should the value chosen for weighting each region contain a large value range, or the base map contains regions that are extreme in size (e.g. in a European context: Luxembourg, and Germany), the result might prove to be too distorted for a meaningful representation.

Non-contiguous cartograms apply the same rule for weighting a chosen value to each region, but the topology in this case is discarded. The general method involves scaling regions proportionately, centred on the region's centroid. We believe that one of the main drawbacks to this method of representation is that it works best if all the regional areas are of a similar size and shape initially; any extremes in size difference will result in a loss of meaningful communication.

The circular cartogram or 'simple-shaped cartogram' (R Inoue, 2011) use abstract shapes to represent each region in a pseudo-contiguous arrangement based on a spatial 'best-fit' scenario. The most popular example to date is the Dorling Circular type, where circles representing each region change in size based on a given value (as circles subjected to absolute scaling), they are then jostled into a 'best-fit' position based on two opposing parameters: a process that repels overlapping circles, and simultaneously attracts each circle to a point of least displacement from its original position (Daniel Dorling, 1996; Ryo Inoue, 2011). It could be argued that the term 'circular cartogram' is a misnomer, as any shape or symbol can be used, for example the rectangle (Raisz, 1934) or the hexagon (Kardos, Benwell, & Moore, 2005).

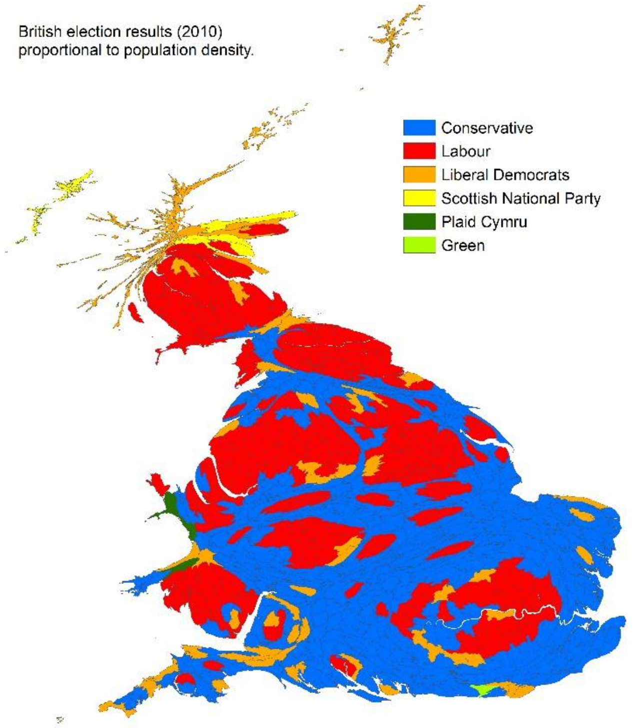


Figure 2: Cartogram showing the 2010 UK Election results.

2 Methodology

Before any two styles of map were chosen it was essential that both map types would have to show almost identical sets of information, allowing a situation where a genuine comparison of communicative readability between the two map types can be achieved. The datasets required for the test should cover the same spatial, and temporal, extent (i.e. records showing one complete year of user activity for an identical area). Two sets of data within the BGS fulfilled this criteria, these were the British Geological Survey's (BGS) iGeology app usage logs (showing activity between October 1st 2010 – October 1st 2011), and the mySoil app usage logs (showing activity between June 1st 2012 – June 1st 2013). The BGS have provided open access to its geological data, via smartphones and tablets using the iGeology app, since September 2010, and the mySoil app since June 2012.

The next dataset necessary for this test is statistics on population distribution. These statistics enable the possibility to quantify app usage alongside the number of people in that area. The Office of National Statistics (ONS) release a national census of Great Britain once a decade. Data collated by this census provides an opportunity for researchers to obtain a snapshot of the numbers of people residing within Great Britain at that time. Choosing a population census nearest to the time the app usage data was created provides a more authentic picture of population usage. The most appropriate release of population census statistics at the time of study was the 2011 census. The next consideration was choosing which scale of boundary type to use in association with the census data. For the test to work, all polygon areas related to the census data, had to be easily viewed, with as little visual bias (differences in region size) as possible.

Creating a base map with a diverse size range of regions could potentially bias the reader towards larger polygon areas. Based on careful consideration between various aggregated levels of boundaries and their associated census statistics, the UK Data Service 'English Westminster Parliamentary Constituencies 2011' boundary dataset, clipped to the OS Boundary-Line 'high_water_polyline', was chosen for the test (Figure 3).

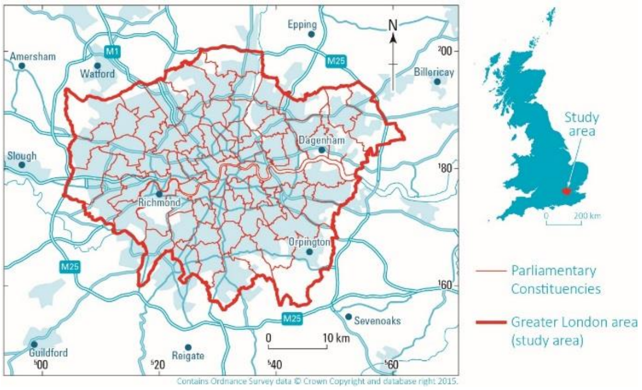


Figure 3: Location of the study area used in the test showing boundaries used for polygons.

2.1 APP USAGE RATIO

One way to effectively 'normalize' population levels using genuine distribution characteristics would be to calculate population proportionally with app tile requests, using an app usage ratio. Such a ratio was created for the test (**Equation 1**), demonstrated in the following: "If, for example, area A has a population of 10, and area B a population of 100, and both areas revealed a tile request total of 10; then area A would have an app usage ratio of 1:1 and area B an app ratio of 1:10; making area A the area with the most app usage." Based on the normalizing characteristics of this ratio, and the perceived ease in reading such a number, this method was chosen to communicate app usage.

Before creating this app usage ratio, certain criteria had to be established. These criteria were followed in order to reduce the level of ecological fallacy, inherent in any generalised statistic. It was noticed through analysis of the app log dataset that every time a user activated the app, the app log recorded the scale of map that was displayed on their screen; these scales are known as 'zoom' scales. For the purpose of the test, the number of tile request were restricted to the 'Zoom 14' scale map tiles, as these were considered to be the most likely scale that indicated intended app use. Tile scales smaller than 'Zoom 14' were considered more likely to reflect lower level activities, e.g. apps that were activated briefly, only to be closed down again; or requests viewed during a manual sweep (search) to a desired area of interest. Next, outlier sums (spikes) in tile requests, deemed either not in the spirit of genuine app usage, or unrepresentative of the main body of the mode frequency of app usage were identified. Records showing less than 2 tile requests for an individual IP address, at one point in time of activity, were removed, as these were considered most likely to represent apps that were activated briefly, only to be closed down again. Tile requests that reached a level far beyond the general statistical group, such as a high number of tile requests made by a single user, were also filtered out; since these would distort the mean average required to make the ratio. Once both incidental and abnormal tile request activity was removed, the mode of tile request was identified. This result was then applied as a reasonable mode marker for the ratio equation. The mode (Mo) for iGeology was 8 tiles, and the mySoil app w 12 tiles. Equation 1 shows the formula used to demonstrate app usage for each map.

$$U\phi = 1 \div \left[\frac{\left(\frac{\sum}{\mathrm{Mo}} \right)}{N} \right]$$

$\sum$ = Number of Zoom 14 (large scale) map tiles requested
Mo = Mode (the 'most frequent' sum of tiles requested)
N = Population (number of people)
$U\phi$ = Usage Ratio

Equation 1: App usage ratio formula.

2.2 THE STUDY AREA

Our initial intention was to apply the new app usage ratio to the chosen ONS Parliamentary Constituencies dataset for the whole of Great Britain. However, the range in polygon size across the constituencies proved too high. Analysis revealed a 99.94% difference in size between the largest constituency of "Ross, Skye and Inverness West" (31,146.2km^2), and the smallest "Islington North" (19.06km^2). It was believed that such diversity in polygon size would draw the reader's eye towards the larger polygons. If this were to be the case, this effect would have been detrimental to the communicative ability of the choropleth map. Furthermore, the cartogram would have benefited from the normalising effect of resizing regions by population (Dykes & Unwin, 1998), further biasing readability in favour of the cartogram. Research by (Kaspar et al., 2011) and (Nusrat & Kobourov, 2015), suggests that cartograms showing extremes in polygon sizes would make it harder for readers to detect magnitude change. Applying the same census statistics to the cartogram, using an exaggeration factor of 10 only reduced the difference in size between the largest and smallest constituencies by 0.19%. Although further exaggeration would have reduced the area difference, it became quite clear that all statistical exaggeration beyond a factor of 10 distorted the map beyond readability. Based on the drawbacks stated above, the focus of this research moved away from full dataset coverage to a more regional study area.

Two factors helped influence our choice of study area. Firstly it was important that knowledge of the area would not influence how participants answer the questions. To address this, it was important to choose a study area that would not be based on a local level such as Nottinghamshire where the majority of participants would likely be familiar. Secondly was that the IDW map didn't communicate effectively where there was deemed to be high app usage. Analysis of the map revealed that many constituencies classed as high app usage actually contained various levels that could show differentiation. An analysis of the 73 Greater London parliamentary constituencies revealed a significant reduction in size ratio between the largest and the smallest polygon areas; 90.55% difference in size between the largest constituency of “Orpington” (201.6km^2), and the smallest “Islington North” (19.06km^2). When applying the same census statistics to the cartogram, using an exaggeration factor of 10, a reduction of 34.19% was measured in size difference between the two constituencies. The extent of the study area is shown in Figure 3.

A decision was made to present the maps in printed format, as opposed to onscreen visuals, in the belief that this manual method would allow participants every chance to read the maps through their own method of adjustment (i.e. distance, orientation, and light conditions); as might be the case in reality, when viewing a journal, magazine, or newspaper. Participants would also be able to mark their answers directly on to the paper map, at the point when they reach an answer to any question presented to them. The test questions, designed to help identify the difference in the map’s communicative effectiveness, would need to be simple, concise, queries. The test should keep to a minimum number of questions, so as not to lose the interest of the participant. Once all participants have completed the test, the results can then be collated and analysed for their communicative potential. The 73 constituencies of the Greater London district appeared to be an ideal candidate for both cartogram and choropleth map use, because all the polygons in that area were of a similar size; reducing, as naturally as possible, the effects of bias towards any large regions and the potential to overlook the smallest regions.

2.3 TEST PROCEDURE

An empirical test was performed using both styles of map to visually demonstrate statistics based on genuine data: mobile device app user activity in the Greater London region of the United Kingdom.

Prior to testing, a procedure was designed that would help educate participants on both map types to be presented to them. This ‘pre-preparation’ exercise involved invigilators providing a short presentation demonstrating how you interpret a choropleth map, and a cartogram, by showing an example of each map.

For the test, each map type used different data to ensure that there was no familiarisation with the data that could influence how the maps performed. This was achieved by splitting participants into four groups. Two sets of groups tested the choropleth map, followed by the cartogram; and two sets tested the cartogram, followed by the choropleth maps. Of the two groups that were testing the same type of map, the data was shown in a different order - one group tested the iGeology data first, followed by the mySoil data, and vice versa for the other two groups. The performance for each map type could then be measured to see whether this affected how maps communicate once a participant is ‘warmed up’ to the test situation through reading the maps. The order that the map types and app usage data was presented to participants is shown in Figure 4.

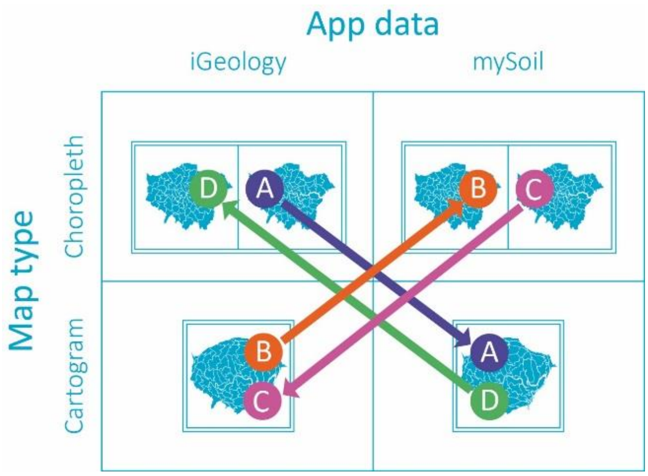


Figure 4: The order that the map types and app usage data was presented to participants.

Each participant performed the test under the same conditions, with two invigilators present. The participant was given a paper document, explaining the test procedure. At the same time, the test invigilators would verbally inform the participant the same information contained within the document. It was believed necessary to introduce each participant to the purpose of the study. Prior to their arrival, their only knowledge of the test was that they had only been asked to participate in a ‘map study’. The introduction helped set the scene as to the nature of the test, and why the BGS was investigating this research area.

The whole procedure was designed to take no longer than 20 minutes, in order to avoid participant test fatigue. The test contained the following elements:

- An introduction to the test and why it was being performed.
- An explanation of a choropleth map and a cartogram.
- A participant consent form.
- A participant information sheet.
- Three questions that relate to one map type.
- Three questions that relate to the other map type.
- A question asking which map type the participant preferred and why.

It was essential that each participant understood all aspects of the test so as to prevent bias due to a lack of participant understanding. To achieve this, an explanation of the two types of map was carried out prior to the test. The difference in how the data is displayed in each map type was made apparent and their technical aspects described. This proved worthwhile, as very few of the participants had even heard of these maps prior to testing; although many of the participants had seen one or both map types, without realising what the map types were called. Once the participant had been given all the information about the test, we asked that they sign a participant consent form detailing that their test results would not be shown as a named individual, and to ensure that the study abided by the BGS research ethics. At this point a participant could choose whether to continue with the study or not. Participants were asked for details on their gender, age, any visual impediments that could affect the answers that they gave, and a self-assessment of their own map reading ability. In order to gauge map reading ability the participant was asked, "How do you rate your own confidence in reading maps?" A scale range between 1 (Low) and 5 (High) was provided. To help the participant judge their own abilities a scenario was provided, stating: "If you are on holiday in another part of Great Britain and you are watching the weather forecast on television; how confident would you be in finding your location on the weather map?" Due to the time and resources available it was felt that this would be the most appropriate way to get an assessment that would reflect how the participant felt about their ability of reading maps.

Other methods of assessment were considered including a pre-test evaluation, e.g. using a map to identify a location. However, this method was dismissed on the basis of prolonging the process too much before the main test.

2.4 TEST QUESTIONS

Before the main questions were given, each participant was shown the maps prior to testing. Each map type was presented to the participant in printed format and its content fully explained, e.g. one choropleth map showing BGS iGeology app usage activity over a one year period, using boundaries based on UK Data Service 'English Westminster Parliamentary Constituencies 2011' boundary dataset, clipped to the OS Boundary-Line 'high_water_polyline' feature layer. The second choropleth map used an identical base map to represent population for each polygon region. Each map was coloured according to the data they represented. One contained app usage (based on a usage to people ratio), classified into ten category ranges; the other contained population level (an aggregated sum of all people within that region as sourced from the 2011 population census data), classified into ten class ranges. The categories on each choropleth map was then explained to the participant along with the requirement to choose one polygon to answer each question, once testing commenced. The maps that were used for each of the questions are shown in Figure 12. Presenting the maps in hard copy format guaranteed that all participants would be viewing the material at the same scale in identical environmental conditions, as opposed to viewing the maps remotely on a computer screen of unknown specification and environmental context.

All participants were asked three questions separately (2.5). The participants were then required to identify and mark just one polygon on the map that they perceived as being the best choice to answer that question. Both choropleth maps had to be used in conjunction with each other to get the most accurate answer. The participant was not given any answers to the questions once they had responded. Similarly the participant could not change their answer once it had been marked down.

The next three questions were all based on the other set of data, and map type, for example, BGS mySoil app usage data using a cartogram. In this case a single cartogram was used to represent the app usage as a ratio in the 10 coloured bin categories and the population level was used to distort the polygons. If the polygon contained a higher than average population proportional to all of the other polygons, it expanded in size. If it contained less than the average population proportional to all of the other polygons, the polygon reduced in size. The only topographical cue, beyond the shape of the polygons in the cartogram, was the polygon boundaries revealing the outline of the River Thames.

All answered questions were measured the same way. Firstly by the statistical accuracy of the answer, given (through choice of polygon); and second by the time taken for the participant to answer the question. Each question was presented to the participant on a separate piece of paper, at which point the clock was initiated by one of the invigilators. Once the answer was clearly marked on the map, the clock was then stopped. The time taken between receiving the question, and marking the map, was recorded by an invigilator.

The final question allowed the participant to choose which map type they preferred, and explain why they preferred that map. This allowed for qualitative information to be collated based on the participants feelings towards how the data was displayed. This was the only opportunity during the test to gauge and record the participant's reaction to the map representations presented to them. The results from this question can determine if the personal opinions of the participant reflect the practical results from the test. In other words, the feelings people have about the maps may not reflect how they performed when trying to interpret the information. This is relevant to record as people's perception of the maps can be just as important as the effectiveness to communicate the information.

2.5 MEASURING EFFECTIVENESS

The aim of the test was to measure the communicative effectiveness of the map types through their ability in communicating the information required to answer the following questions:

1) Out of the highest populated areas, which area had the least app use?
2) Out of the lowest populated areas, which area had the most app use?
3) Which region performed best overall?

Question one and two are considered to be more 'complex' due to the perceived dichotomy of direction between the superlative adverb and the adverb (i.e. highest, and the least; lowest, and the most). In other words, the first two questions are asking for a result that requires the participant to combine opposing (extremes) in values of app usage. In contrast to question one and two, the third question holds no such dichotomy, and is considered to be straight forward 'simple/basic' in its request. These definitions were used by (Kaspar et al., 2011) to some effect, identifying variations in the level of communicative ability between similar map types.

It was decided that absolute answers relating to a single polygon area would not be an appropriate measure of success, or failure, of readability. An absolute answer would favour the cartogram, due to the fact that in the creation of the cartogram absolute values are used to distort the shape of the polygon, whereas the choropleth map aggregates the data into generalised grouped ranges, represented by colour. The aim of this test was to focus on whether participants could detect an overall impression of what the app usage data was showing, rather than picking the absolute answer for each question. In order to measure the effectiveness, the five most correct answers for each data type were calculated. The five correct polygons represented 7% of the possible answers on the map, and was felt by the authors to be a reasonable measurement to determine whether the map type is communicating effectively.

The methodology for determining these answers to each question are as follows:

a) **Out of the highest populated areas, which area had the least app use?**

Highest populated = The population for each of the 73 polygons were numerically ordered so that the highest populated area was attributed with a score of '73'; and the lowest a score of '1'.

Least app use = The app usage ratio for all the polygons were ranked, with the best ratio of app usage ranked as '1' and the worst ratio of app use with a score of '73'.

The two scores were then multiplied together for each polygon, leaving the polygon with the highest value representing the most favoured answer. The five highest results were classed as correct answers.

b) **Out of the lowest populated areas, which area had the most app use?**

Lowest populated = The population for each of the 73 polygons were ranked in order so that the highest populated area was attributed with a score of '1'; and the lowest populated area ranked with a score of '73'.

Most app use = The app usage ratio for all the polygons were numerically ordered, with the best ratio of app usage scoring '73', and the worst ratio of app use scored with a '1'.

The two scores were then multiplied together for each polygon, leaving the polygon with the highest value representing the most favoured answer. The five highest results were classed as correct answers.

c) **Which region performed best overall?**

The app usage ratio used for this study (Equation 1) only represents a proportion of app usage as a fraction; there is no quantitative value on its own. Only by using the fraction consequent (q), with a real value, can a quantitative result be achieved. In order to reveal the best performing region, the true value of the fraction consequent needs to be used to divide up the population count for that region. For example, two areas with an app usage ratio of 1:10 only demonstrate that one in ten people used the app in those regions. However, if you use the consequent number (q = 10) to divide up the number of people in that region, a definitive answer can be achieved. So, if 10 people lived in region A, and 1000 people lived in region B, region B demonstrates the highest overall level of app usage, with a result of 100 people. Therefore, by using population level divided by the ratio of app use, it was then possible to identify the top five performing regions.

The answers to these questions would reveal how accurate the participants performed for each of the map types. Other questions relating to the participant's age, gender, visual impediments, and map preference, determine if any of these factors had any effect on participant performance for both map types.

3 Results

A total of 72 participants completed the test; 36 male and 36 female (Holbrook & Cartwright, 2016).

3.1 CORRECT ANSWERS

Initial focus aimed at establishing how many participants answered questions correctly for each map type presented. A correct answer was considered to be one of the resulting top five polygons identified using the methodology above.

A broad overview of which map type performed best shows that out of a total of 216 questions asked for each map type, 146 were correctly answered for the choropleth map, and 154 for the cartogram. This translates to 68% of participants answered the questions correctly for the choropleth map, and 71% of participants answered the questions correctly for the cartogram; a 3% margin of difference between both map types, therefore not a conclusive result in support of one map type being more effective at communicating than the other.

3.2 MAP SEQUENCE

Griffin (1983) noted that participants adapted to the cartogram as they gained experience in reading them. Inspired by this observation, we decided to explore if there was any influence on participant ability to read the maps based on the map type sequence. Initially we looked at the how the participants performed on their first (respective) map type presented in Figure 5. The aim was to establish if participant performance was either enhanced, or impeded, on their reading of the first map in the test. Did the participant need to 'warm up' in order to interpret the data effectively?

Two of the participant groups looked at the choropleth maps first. Therefore, 36 participants responded to questions based on the choropleth maps before moving onto the cartogram; and vice versa. 108 questions were asked relating to each map type. For the choropleth, 73 questions were answered correctly (68%). For the cartogram, 78 questions were answered correctly (72%); showing an overall difference of 4% between the two map types. However, analysis of the results in the context of the three individual map questions, revealed a more complicated picture. As described in the previous section, the first and second question, relating to the map, were classed as 'complex' questions, and the third was classed as the 'simple' question. For the first question, 26 out of the 36 participants, in their respective groups, read the choropleth correctly (72%). For the cartogram 27 participants out of the 36 read the map correctly (75%). The second question resulted in 28 out of 36 participants read the choropleth correctly (78%); the cartogram, 23 participants (64%). The third 'simple' question resulted in 28 questions answered correctly for the cartogram (78%), whereas the choropleth map returned 19 questions answered correctly (53%).

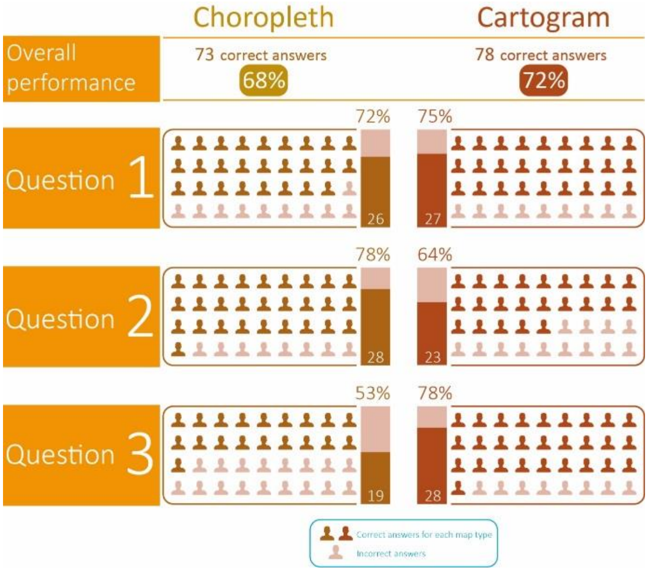


Figure 5: Performance of each map type, based on the first map tested by the participants in the test.

The process was repeated for the second map (Figure 6) showing the following results: For the choropleth, 77 questions out of the 108 were answered correctly (71%). For the cartogram, 75 questions were answered correctly (69%); showing an overall difference of only 2% between the two map types. Further analysis of the results in the context of the three individual map reading questions revealed, once again, a more detailed picture. For the first question, 22 out of the 36 participants, in their respective groups, read the choropleth correctly (61%). For the cartogram 26 participants out of the 36 read the map correctly (72%). The second question resulted in 31 out of 36 participants read the choropleth correctly (86%); the cartogram, 21 participants (58%); a difference in favour of the choropleth maps by 28%. The third 'simple' question resulted in 28 questions answered correctly for the cartogram (78%), whereas the choropleth map only returned 18 questions answered correctly (50%); showing a 28% difference in favour of the cartogram.

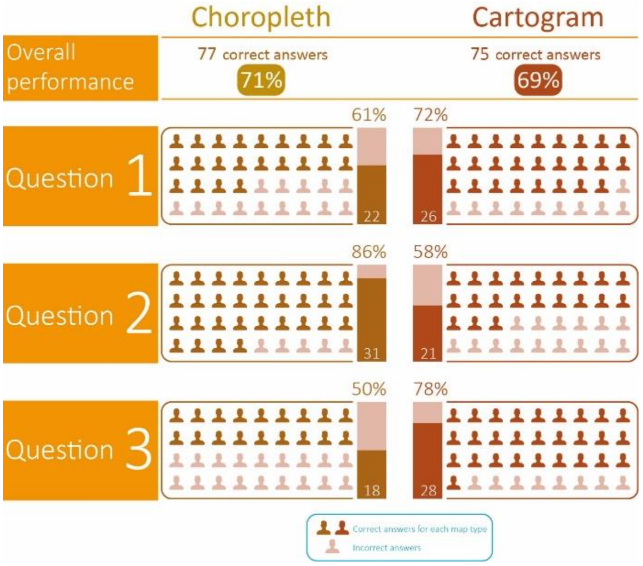


Figure 6: Performance of each map type, based on the second map tested by the participants in the test.

Overall, results demonstrate that the map sequence did not significantly influence participant performance during the test.

3.3 DATA SEQUENCE

Results were analysed for differences in participant performance based on the order by which the data was viewed; this was to establish if viewing sequence of the two types of app usage data (iGeology app and mySoil app usage) influenced how people read the maps. All participants saw both sets of data, albeit in either cartogram or choropleth map form. 108 questions were asked relating to each data type. For the first round of maps presented to the participant (Figure 7), 85 questions (79%) were answered correctly for iGeology, and 66 questions (61%) were answered correctly for the mySoil app usage data; showing an overall difference of 18% between the two sets of data. For the second round of maps, 74 questions (69%) were answered correctly for iGeology, and 72 questions (67%) were answered correctly for the mySoil app usage data; showing an overall difference of 2% between the two sets of data.

Breaking this down to individual question level revealed the following for the first round of maps presented to the participant: The first 'complex' question revealed 30 out of the 36 participants read the iGeology map correctly (83%), and 23 out of the 36 participants read the mySoil map correctly (64%). The second 'complex' question revealed 26 out of the 36 participants read the iGeology map correctly (72%), and 25 out of the 36 participants read the mySoil map correctly (69%). Question three 'simple' revealed the greatest difference with 29 out of the 36 participants (81%) reading the iGeology map correctly, with only 18 out of 36 participants (50%) reading the mySoil maps correctly.

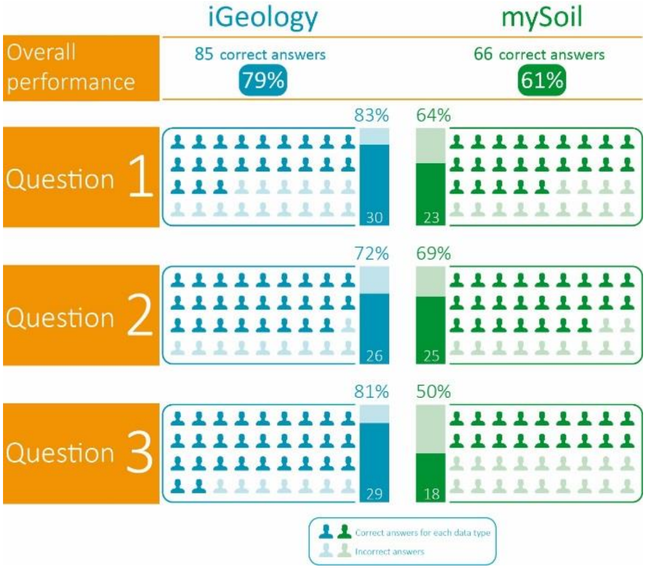


Figure 7: Performance of app usage data, based on the first map tested by the participants in the test irrespective of map type.

For the second round of maps (Figure 8): The first 'complex' question revealed 27 out of the 36 participants read the mySoil map correctly (75%), with only 21 out of the 36 participants reading the iGeology map correctly (58%). The second 'complex' question revealed a 72% success rate (26 out of the 36 participants) for both iGeology and mySoil maps. Question three 'simple' revealed the greatest difference with 27 out of the 36 participants (75%) reading the iGeology map correctly, with only 19 out of 36 participants (53%) reading the mySoil maps correctly.

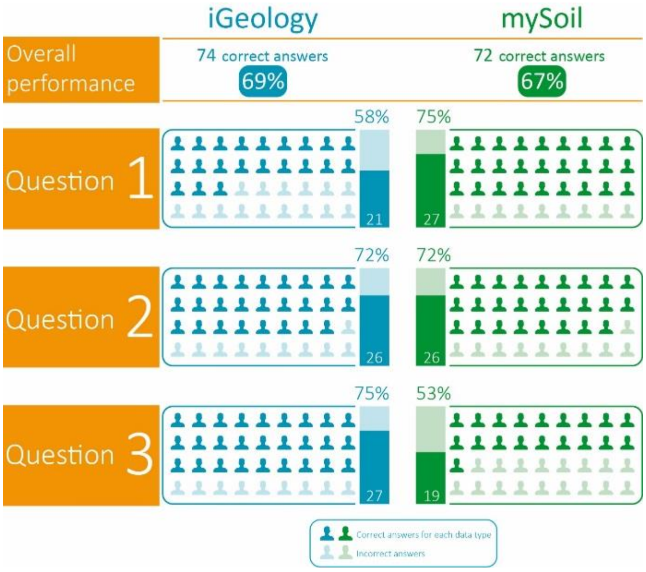


Figure 8: Performance of app usage data based on the second map tested by the participants in the test irrespective of map type.

These results show that the maps containing iGeology data performed better overall. Evidence revealed a poorer performance by participants reading mySoil data in the first round but the trends were relatively consistent for the second and third question, regardless of the order the data was seen in.

3.4 THE QUESTIONS

Results were analysed for differences in participant performance for each map question, with an emphasis on the complexity of each question. Question one and two are considered to be more complex due to the perceived dichotomy of direction between the superlative adverb and the adverb (i.e. highest, and the least; lowest, and the most). In contrast to question one and two, the third question held no such dichotomy, and is considered to be straight forward 'simple' in its request. The results from this aspect of the study are shown below.

Question 1 was: Out of the highest populated areas which area had the least app use?

Of the 72 questions asked relating to each map type 48 (67%) were answered correctly for the choropleth map, and 54 (75%) correct answers for the cartogram; a difference of 8% in favour of the cartogram.

Question 2 continues the 'complex' questions, this time asking: Out of the lowest populated areas which area had the most app use?

This time the choropleth maps were more successful with 59 questions answered correctly (82%) compared with 44 out of 72 (61%) answered correctly for the cartogram; a difference of 21% in favour of the choropleth map.

Question 3 is the 'simple' question asking: Which region performed best overall?

Of the 72 questions asked relating to each map type 56 (78%) were answered correctly for the cartogram, and only 37 (51%) correct answers for the choropleth map; a significant difference of 27% in favour of the cartogram.

These results highlight a clear difference in map performance between the two types of question used. The results reveal the choropleth map performing better with second 'complex' question and the cartogram proving more effective at communicating the information necessary to answer question 3.

3.5 AGE OF PARTICIPANTS

The age of the participant was also analysed to see if there was any apparent effects between the legibility of the two map types (Figure 9). Of the 72 participants, there were four age group categories: 25-34, 35-44, 45-54 and 55-64. There were variations as to how many participants were represented by each group and these were as follows:

25-34	14 participants
35-44	22 participants
45-54	26 participants
55-64	10 participants

For question 1 there were no significant differences in how the map types performed. Overall the cartogram had the edge in being slightly more successful. There was a maximum of 8% difference between the performance and this was on the 35-44 category with the cartogram achieving

68% of participants giving the correct answer and 59% for the choropleth map. The choropleth maps only achieved a better result in the 45-54 category but this was just by 4%.

Question 2 shows a switch in performance in favour of the choropleth maps. Nearly all the age group categories perform better on the choropleth map, apart from the 55-64 age group which performed equally on both. The difference reaches as much as 38% in the 45-54 category with 88% answering correctly when using the choropleth maps, compared to just 50% of participants being able to answer correctly using the cartogram. A difference of 18% is also apparent in the effectiveness of the choropleth map for the 35-44 age group, and a reduced margin of 7% for the 25-34 category.

Question 3 shows a complete reversal of effectiveness for the choropleth map. This time the cartogram performs better across the board for all age groups. On average the difference between the two map types is 27%. The 25-34 age group has the most difference with 36% correct answers through reading the cartogram; followed by 35% for the 45-54 group, 20% for the 55-64 group, and 18% for the 35-44 age group.

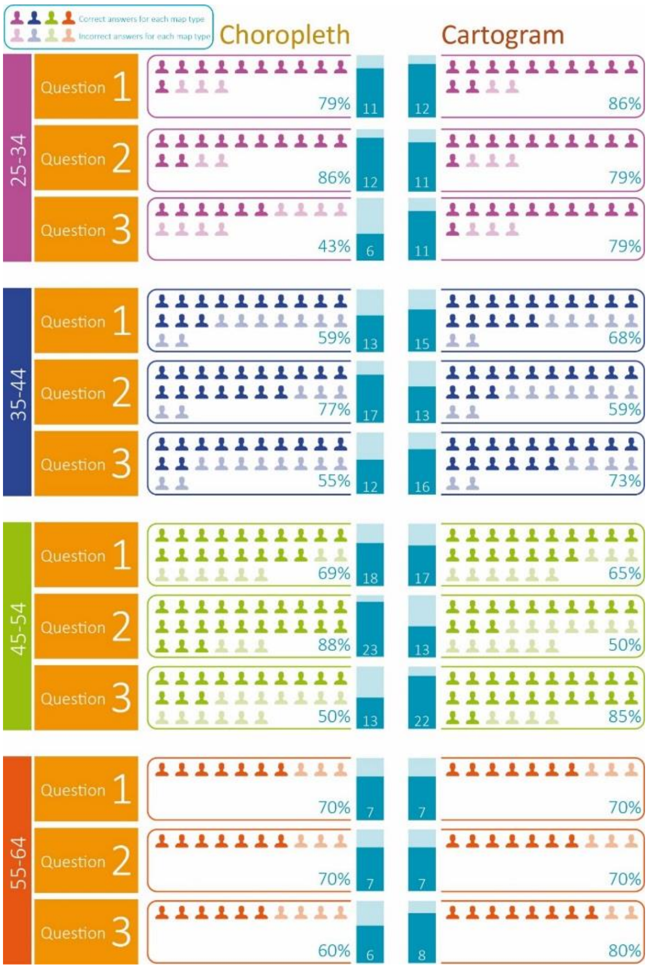


Figure 9: Answers given by participants for each question based on map type for each age group.

3.6 TIME OF RESPONSE

All participant results were timed, adding a temporal measure of performance for all map questions (Figure 10). Overall, participants read the cartogram on average 5 seconds faster than the choropleth map. However, in the context of the age group categories, the 45-54 age group read the choropleth map almost 4 seconds faster than the cartogram, with a 2% advance on the accuracy of their answers. The other groups all performed better with the cartogram and were quicker (with the exception of the 35-44 age group who were slightly slower). The most significant result appeared from the 55-64 age group, where 80% answered questions correctly from the cartogram (compared with 63% for the choropleth maps), on average 11 seconds faster. The 25-34 age group showed 81% of participants reading the cartogram correctly compared with 69% for the choropleth maps, 5 seconds faster than the choropleth maps.

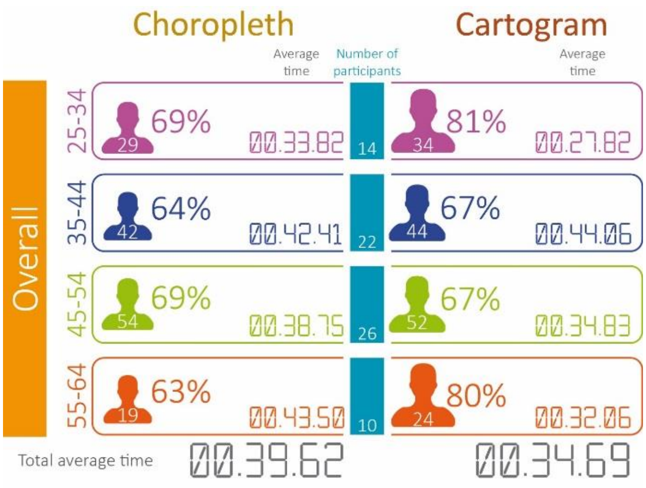


Figure 10: Performance of each map type based on the age of the participant including response times of correct answers.

3.7 MAP PREFERENCE

At the end of the test, each participant was asked which map they preferred. Of all the participants, 75% chose the choropleth map as being more preferable; with the main reason being ‘familiarity’. With this result, we compared how each map type performed against the participant preference (Figure 11). Overall performance shows that for those participants that preferred the cartogram, they performed almost consistently on both types of maps (71% correct answers for the cartogram and 70% for the choropleth map). From those that answered the question correctly, the cartogram was more efficient, and enabled the participant to deliver an answer 1.77 seconds faster than the choropleth map.

The results from the participants that preferred the choropleth map revealed a significant difference in performance between the two map types. Even though the choropleth map was preferred, the participants performed better with the cartogram with 64% giving a correct answer, compared to 60% with the choropleth maps; with the cartogram enabling the participant to deliver an answer on average 6.43 seconds faster than the choropleth map.

Results were also analysed for differences in participant performance for each map question. Despite the participants’ preference of map, the cartogram performed better in correct answers in response to the first question, by an average of 8.5 seconds. Of the participants that preferred the cartogram, 78% gave correct answers from the cartogram, compared to 62% for the choropleth maps. The ones who preferred the choropleth map answered 66% of questions correctly from the cartogram, but only managed 54% correct answers using the choropleth map.

Question 2 shows a different picture. Of the participants that preferred the choropleth maps, 74% managed to answer the question correctly, compared to just 54% with the cartogram; albeit on average 8.87 seconds longer to read the choropleth maps. For the participants who preferred the cartogram, 62% managed to answer the question correctly, compared to 86% using the choropleth map. Although the choropleth map took an average 13.06 seconds longer to reach a correct answer, it was 24% more effective than the cartogram.

Question 3 reveals a significant difference in performance between the preferred map types. Of the participants that preferred the cartogram 72% managed to answer the question correctly, compared to 62% with the choropleth maps. There was only a 1.19 second average improvement in efficiency for the cartogram. However the participants who preferred the choropleth maps actually performed better with the cartogram for question 3, with 74% giving a correct answer, and 52% for the choropleth maps. The time it took the participants to provide a correct answer for the choropleth maps was significantly slower. The choropleth maps took on average 18.17 seconds longer to read, compared with the cartogram.

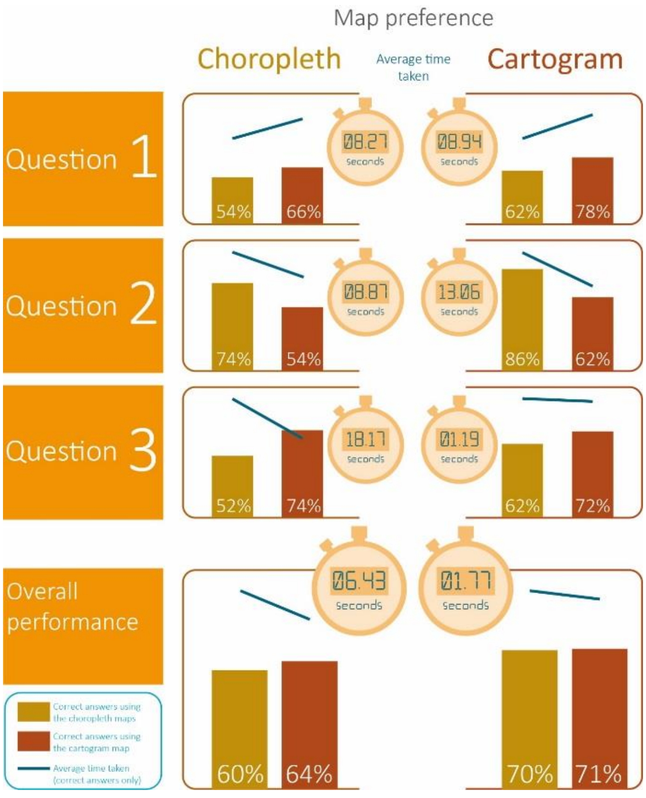


Figure 11: Performance of each map type based on participants preference of map type. Average time taken for correct answers given is also shown.

3.8 OTHER RESULTS

Results were analysed for any variations of communicative effectiveness based on the sequence of map type presented to the participant during the test. The results did not show any difference of note, but there is evidence that the performance of some participants may have been slightly impaired on the first question of the cartogram, as observations during the test witnessed a 'warming up' process by some participants.

Results based on what participants thought of their own map reading ability revealed the cartogram to be easier and quicker to read for those that classed themselves to have either, good or, excellent map reading ability. Those that generally declared themselves as average or less in map reading ability managed to read the choropleth maps more effectively.

Overall, the results show that female participants found the communicative effectiveness of both types of map to be the same, with 62% gaining correct answers from the cartogram, and 64% from the choropleth maps. The male participants, however, found the cartogram better to read with 70% correct answers, and 61% correct answers as result of reading the choropleth maps.

Results based on what participants thought of their own ocular ability remain inconclusive, as only 3 out of the 72 participants declared that they had a minor impairment. It was decided that this was an insufficient number of participants to make any meaningful assessment of measurement.

4 Discussion

Further analysis of the results have produced a number of significant findings, both new and in support of previous published work. Our observations and some insights are discussed below.

4.1 COMMUNICATING MAGNITUDE

One of the key aspects of this test was to establish to what extent people were able to read magnitude from a cartogram, without the supportive 'linking' aids used in previous research (Dent, 1975; Griffin, 1983; Kaspar et al., 2011; Nusrat et al., 2015; Sun & Li, 2010). In recent years, it has been noticed that the practice of using a magnitude key alongside an inset map appears to be largely overlooked in published material. Based on this observation, we wanted to establish if the use of a cartogram without such an aid was prudent. The results from the test show that communicative effectiveness of the cartogram is not hindered by the lack of an inset map; on the contrary, the cartogram generally gains over the choropleth map in both accuracy and communicative efficiency. Only in the context of question 2 does the choropleth gain the advantage over the cartogram, in communicative accuracy. For question 2 participants were required to identify high app usage, which was displayed through the use of colour, against the lowest populated areas; magnitude visualised through the shrunken polygon areas.

During the test it was clear that participants found low populated areas difficult to identify, and in some cases, the presence of shrunken polygons were not acknowledged. In these cases, the participant appeared naturally drawn towards the larger polygons (displaying higher app usage), showing that size was the greater influence in the process of choosing a region. However, the cartogram did manage to convey the correct answer more successfully for questions 1 and 3, where the participant was looking for highly populated areas, meaning that larger polygons had to be identified. The results show the cartogram is good at communicating statistical data that expands polygon areas, as the eye is not only naturally drawn to these areas, if the conditions are correct, they are better at communicating magnitude. The majority of participants successfully gained a relative scale of magnitude from the cartogram, without access to a projected inset map. Why this should be is not known, but there is one visual cue that may be the key towards the mechanism behind the brain judging magnitude so successfully. All the polygon areas used in the maps are based on parliamentary constituency boundaries, created historically via influence along both anthropogenic and geographical features on the ground. The map reader viewing these 'generalised' boundary lines is able to evaluate relative magnitude from the unnatural series of directionally skewed, acute angles, created by the distortion making process of the cartogram. A study by Kaspar et al., (2011) noticed that irregular polygon boundaries communicated magnitude better than regular shaped polygons. By looking at the area, and distorted boundary edges, the mind appears equipped in sensing an accurate change in magnitude without pre-distorted knowledge of the area. The test results support the notion that these distorted polygons contain sufficient information for the brain to quantify a measure of relative dynamic magnitude, without the aid of a choropleth map. Further study in this area may establish if this is the case, and help to identify what thresholds of effectiveness exist e.g. the effective limits of the exaggeration factor.

4.2 THE LINKING METHOD

Unlike the cartogram, the use of two statistical datasets meant that it was appropriate to use two choropleth maps, rather than a single choropleth map using a combination of colour and patterned ornaments. By separating the two datasets into two choropleth maps, this avoided the potential of obscuring the map with new features (distracting, overlaying visual clutter), not present on the cartogram. It was thought that additional information augmented onto the choropleth maps would have made the two map types visually inconsistent, creating an unequal comparison of communicative effectiveness. As a result, the separation of the datasets into two maps not only enabled aesthetic consistency, it also provided an opportunity to observe the effectiveness of maps using the linking method of communication. Both choropleth maps used the same size, scale, polygon boundaries, and key colour ramp; with only the data in the key content to distinguish between the two maps. During the test it was noted that the comparison between the maps sometimes caused a problem in identifying the same polygon areas on both choropleth maps. On occasion, a participant would identify a polygon on one map and believe they had identified the same polygon on the second map, which wasn't actually the case. This meant that the answer given would be inaccurate and the data wasn't used correctly from both maps. Furthermore, the results generally showed that the time taken to reach a decision between the two choropleth maps was significantly longer than the time taken to read the cartogram.

4.3 COLOUR RAMPS

The colours used on the maps were specifically chosen to cater for colour blindness. These colours were selected via choices presented on the ColorBrewer2 website. The participant was asked prior to the test if they had any visual impediments, meaning that even if they were colour blind, it should not have had a negative effect on how each map type performed. The number of bin values used was limited to ten, as the perception of colour is far more difficult when relating to a legend above this amount (Dent, 1990). It was found that there was generally not a problem with relating a colour to the legend on the maps, but the colours themselves could attract a participant to choose a polygon purely based on colour. Participants were often drawn towards the darker, more dominating colours; betraying a degree of colour bias, influencing their choice of answer. Colour ranges used on both map types can lead to a misinterpretation of the data, with participants choosing polygons that they thought were the same on both maps relating to colour, rather than statistical accuracy. It was observed that a habitual interpretation of certain colour ramps can override the message being communicated by the map. For example, people that read maps using a certain colour ramp on a regular basis find it very hard to re-orientate the range polarisation from one direction to another. For example, a hydrologist will read a red to blue colour ramp in the map key as 'dry to wet' (negative to positive); however, apply this to a map that states the same colour range as 'high to low' app usage (positive to negative), confusion may develop. One way to avoid this would be to identify your audience and keep from using any colour key convention the reader may be conditioned to. However, if the map is to be aimed at a wider audience, try to avoid the common combinations such as: Red to Blue. Further analysis into colour perception would help identify a colour scheme that would be more objective and perhaps appropriate to the subject of your statistics, preventing the pitfalls of a predisposed influence to a particular colour range.

4.4 QUESTION COMPLEXITY

Results were analysed for differences in participant performance for each map question. Question one and two were labelled as 'complex questions' due to the perceived dichotomy of quantitative direction between the superlative adverb and the adverb (i.e. questions asking for the highest, and the least; or the lowest, and the most). Question three was labelled as a 'simple question', due to there being no dichotomy of quantitative direction in the question (i.e. highest, and the most). The results show the choropleth to be more successful at answering the complex questions. Question two especially revealed how successful the choropleth map was at communicating these statistics, where the participant is looking for extremes from opposing levels: population and app usage. This result supports previous work which identified that choropleth maps communicated complex questions more effectively than the cartogram

(Kaspar et al., 2011). So, the more complex the question, the message was received more successfully by using choropleth maps.

The benefits to using a cartogram was demonstrated rather dramatically in question three, where it was far more effective at communicating the appropriate information than the choropleth maps. The cartogram performed better overall in all four of the participant groups, and was significantly quicker to read. This was the case regardless of the type of data used, or the order that the maps were presented to the participant. These results demonstrate clearly that cartograms should be used to answer simple questions, whereas the choropleth maps are more appropriate when the question is more complex.

4.5 PERSONAL PREFERENCE

Results were analysed for any contrast relating to participants map preference and their performance. Only 25% of participant's preferred the cartogram; based on the qualitative statements written by each participant at the end of the test. The dominant reason behind this 25% was that they thought it was easier to interpret one map containing all of the information. The participants that preferred the choropleth map admitted that they could see the merits of the cartogram, but preferred the familiarity of the choropleth maps. Many participants in the choropleth camp believed they would feel more comfortable using cartograms if they had more experience in reading them. The results show that out of the 75% of participants that preferred the choropleth map read the cartogram more successfully than the choropleth map; and on an average 6.43 seconds quicker. However, it should be noted that the communicative effectiveness measured in this test is based on a captive audience; as one would be if reading a peer review paper, or a report relevant to your needs. These maps might not achieve half the success observed in this test if they were subject to the discretion of a passing audience, e.g. in the context of a poster, or a leaflet. There is no data as of yet showing how such maps perform in a casual viewing context. It is quite likely that one map type might be more visually attractive than the other. This initial hurdle is an important one, as the communicative effectiveness of either map type is of little merit if people are not urged to look at them in the first place.

4.6 TEST OBSERVATIONS

One advantage to holding the test in person is that the authors were able to observe how each participant manoeuvred through each question, in order to reach their answers; both by their actions and their verbal responses. Albeit a very labour intensive method, compared to an online delivery, these observations contributed significantly to our understanding of the results, and helped the authors identify various issues and recommendations that would have otherwise gone unnoticed.

4.7 PRE-TEST INTRODUCTION

It was observed that the pre-test introduction may have impeded the accuracy of some answers given by the participant. Part of the pre-test introduction involved showing the participant a choropleth map of the 2010 general election results. The same data was then displayed in a contiguous area cartogram, demonstrating the normalising effects of using population statistics; showing how the less densely populated areas shrank, and the more densely populated areas expanded (most notably, the Greater London area). However, when the participants were presented with the first question (looking at the cartogram showing Greater London constituencies), a small minority looking for the lowest populated areas did not expect to see the lowest areas to be shrunken on the cartogram. It transpired that the confusion was down to the participant remembering the pre-test introduction election result cartogram of Britain, showing all the constituencies of the Greater London area increased in size. The participant tended to learn of their mistake by the second question when they realised that the polygons reacted proportionally to the population levels of the study area, not the whole country. Unfortunately, the conditions of the test dictated that no previous answers to questions could be changed once the participant had made their choice. This meant that even if the participant realised that a mistake had been made, it could not be recorded that they were aware of their mistake, and that they initially misunderstood the cartogram. This may account for a reduction in cartogram success for the first question. But it is an observation in support of earlier studies that people tend to understand cartograms more through experience (Griffin, 1983).

4.8 USING RATIOS

The ratios that we chose to use on both maps were designed to help the end user digest the information more readily, than the raw statistics used in the past. The belief was that the ratio would work seamlessly with the population statistics, and would facilitate the answering process without bias or confusion. In practice, it was observed that the ratio confused some participants. Ratios using larger numbers were often misinterpreted as being higher levels of app usage i.e. larger ratio number was seen as a higher usage number; an observation noted by Spiegelhalter et al., (2011). The use of the ratio was explained to each participant in the pre-test introduction, but some participants did acknowledge that the ratios were confusing, and suggested that they would rather use a scale that included simple statements such as 'high app usage' and 'low app usage'. Furthermore, one could say that the use of numbers in this context was perhaps not necessary; as the ratio was divided into bins using Natural Jenks, and the ten categories created confusion by displaying numbers that were difficult to understand.

4.9 TIME TAKEN TO ANSWER

One further observation of note was that those participants whom tended to take longer than average to reach a decision, usually answered incorrectly.

5 Post-test analysis

On reflection, overall results from the test have provided useful insights into how both cartograms and choropleth maps work. One is related to the level of distortion required to express a desired message from a cartogram.

5.1 EXAGGERATED DISTORTION

There appears to be a general belief that cartograms should be exaggerated as much as possible in order to be a 'real' cartogram. This method of rendering might be based on a visualisation tactic designed to catch the attention of the reader. However, we noticed that over exaggeration of polygon areas tends to bias results to the larger areas ('greater than' average statistic); effectively losing the 'less than' average half of the statistic. If all you wish to communicate is the 'greater than' average results, then sensationally enhancing bias towards these areas is an effective solution. However, only communicating the top end of the statistic might not be the desired result. This paper demonstrates that a careful adjustment of the exaggeration factor, appropriate to both the statistics and the messages you may wish to convey, is not only necessary, but essential. The distortion factor chosen for the iGeology and the mySoil app usage cartograms required a discrete level of exaggeration in order to allow all the polygons to be visible, and readable. This was to allow the participant a realistic chance to identify a level of magnitude from shrunken polygon areas, when a question demanded it. Our results show that when the distortion factor is used subtly, the communicative effectiveness of the cartogram remains effective.

5.2 PERILS OF TACIT KNOWLEDGE

One of the benefits of this report is that it is finding an effective method of communicating mobile app usage, on a static map format, using genuine data. Unlike manufactured data housed in an artificial spatial context, the data used in this paper is authentic. However, using real statistics in a real spatial context does have its problems. All labelled references to places and features were deliberately omitted in order to help prevent any participant tacit knowledge interfering with the reading of the maps. It was essential that the data used on the maps was read in its own right, for us to assess the success or failure of the map types in delivering their information. An added measure to help reduce the effects of participant knowledge would have been to rotate the map 180° (Danny Dorling, 2015). Viewing the map, upside-down, would have reduced the ease of polygon identification, making any identification of areas harder to achieve. The authors failed to realise this tactic at the time of the test. However, the presence of perhaps one of the most recognisable features within the British Isles (The River Thames), would in this case, prohibit the use of presenting the map upside down, as the area and shape on both map formats was too significant for it not to be easily identified. It would not have been beyond the participant to physically turn the paper maps back round, ultimately distorting the answer finding process. Should the area chosen be devoid of any topographic features, contained within a facility making it impossible to rotate the maps back round, the tactic of displaying the maps upside down would have helped significantly in reducing the invasive effects of participant's tacit knowledge of an area.

6 Conclusions

This report aims to establish which map type helps communicate app usage data more effectively. The study has not revealed a single map type that should be used exclusively for all data types. On the contrary, the results gained from this study clearly establish certain rules to consider when choosing the best way to communicate your spatial data.

Generally, the cartogram has the edge over the choropleth map, both in communicative ability and efficiency in delivery (the speed taken to read), especially if the intended question being asked from it is a straight forward one e.g. which area has the largest output? However, choropleth maps are more successful at simultaneously communicating opposing (extremes) in values, for example, messages that relate to data that is less than the average, or negative in value. The most effective way to use a cartogram is to exploit its major asset of areal visual bias, mainly through expanding polygons proportionally by a given value. The results show that the reader tends to gain a good understanding of relative magnitude from expanded polygon areas, especially if the original base map contains polygon areas of a similar size with irregular shaped boundaries. As a result, cartograms of this nature do not require an inset map in order for them to communicate magnitude effectively. Excessive distortion of cartogram polygons is not necessary. Albeit an effective method of gaining the reader's attention, it is most likely that any meaningful representation of magnitude from the shrunken polygons will become unreadable. It is important that only subtle use of exaggeration is applied if it is necessary to communicate both expanded and shrunken values to the reader.

Appendix 1

Answers to each question

Correct answers to question 1

	Cartogram A	Cartogram B	Cartogram C	Cartogram D	Cartogram Total		Choropleth A	Choropleth B	Choropleth C	Choropleth D	Choropleth Total	
Correct answers in the top 5	15	15	12	12	54	75%	15	14	11	10	50	69%
Correct answer average time	00:42.89	00:43.55	00:45.72	00:38.40	00:42.64		00:33.08	00:36.56	00:36.09	00:32.91	00:34.66	
Incorrect answer average time	00:53.88	00:33.60	00:38.34	00:40.68	00:41.62		00:57.96	00:27.18	00:36.12	00:35.17	00:39.11	

Correct answers to question 2

	A	B	C	D	Total		A	B	C	D	Total	
Correct answers in the top 5	11	12	10	11	44	61%	14	15	14	16	59	82%
Correct answer average time	00:42.43	00:31.99	00:26.94	00:31.02	00:33.10		00:45.71	00:41.06	00:44.67	00:40.62	00:43.01	
Incorrect answer average time	00:41.15	00:47.16	00:27.02	00:26.75	00:35.52		00:47.33	00:43.65	00:44.39	01:15.73	00:52.78	

Correct answers to question 3

	A	B	C	D	Total		A	B	C	D	Total	
Correct answers in the top 5	12	14	16	14	56	78%	15	7	4	11	37	51%
Correct answer average time	00:45.57	00:41.85	00:32.18	00:38.69	00:39.57		01:05.17	00:45.16	01:00.81	00:44.44	00:53.89	
Incorrect answer average time	01:10.13	00:58.14	00:54.38	00:35.02	00:54.42		00:44.81	00:47.54	00:32.35	00:43.04	00:41.93	

Overall correct answers

	A	B	C	D	Total		A	B	C	D	Total	
Correct answers in the top 5	38	41	38	37	154	71%	44	36	29	37	146	68%
Correct answer average time	00:43.63	00:39.13	00:34.95	00:36.04	00:38.44		00:47.99	00:40.92	00:47.19	00:39.32	00:43.86	
Incorrect answer average time	00:55.05	00:46.30	00:39.91	00:34.15	00:43.85		00:50.03	00:39.46	00:37.62	00:51.31	00:44.61	

Map sequence

Cartogram first (groups B and D)

Correct answers	B	D		Correct Answer Average time B	Correct Answer Average time D	Incorrect Answer Average time B	Incorrect Answer Average time D
Question 1	15	12	75%	00:43.55	00:38.40	00:33.60	00:40.68
Question 2	12	11	64%	00:31.99	00:31.02	00:47.16	00:26.75
Question 3	14	14	78%	00:41.85	00:38.69	00:58.14	00:35.02
Overall	41	37	72%	00:39.13	00:36.04	00:46.30	00:34.15

Cartogram second (groups A and C)

A	C		Correct Answer Average time A	Correct Answer Average time C	Incorrect Answer Average time A	Incorrect Answer Average time C
14	12	72%	00:42.89	00:45.72	00:53.88	00:38.34
11	10	58%	00:42.43	00:26.94	00:41.15	00:27.02
12	16	78%	00:45.57	00:32.18	01:10.13	00:54.38
37	38	69%	00:43.63	00:34.95	00:55.05	00:39.91

Choropleth first (groups A and C)

Correct answers	A	C		Correct Answer Average time A	Correct Answer Average time C	Incorrect Answer Average time A	Incorrect Answer Average time C
Question 1	15	11	72%	00:33.08	00:36.09	00:57.96	00:36.12
Question 2	14	14	78%	00:45.71	00:44.67	00:47.33	00:44.39
Question 3	15	4	53%	01:05.17	01:00.81	00:44.81	00:32.35
Overall	44	29	68%	00:47.99	00:47.19	00:50.03	00:37.62

Choropleth second (groups B and D)

B	D		Correct Answer Average time B	Correct Answer Average time D	Incorrect Answer Average time B	Incorrect Answer Average time D
13	9	61%	00:36.56	00:32.91	00:27.18	00:35.17
15	16	86%	00:41.06	00:40.62	00:43.65	01:15.73
7	11	50%	00:45.16	00:44.44	00:47.54	00:43.04
41	36	71%	00:40.92	00:39.32	00:39.46	00:51.31

Data sequence

iGeology first (groups A and B)

Correct answers	A	B		Correct Answer Average time A	Correct Answer Average time B	Incorrect Answer Average time A	Incorrect Answer Average time B
Question 1	15	15	83%	00:28.11	00:45.54	00:57.96	00:33.60
Question 2	14	12	72%	00:45.25	00:24.41	00:47.33	00:47.16
Question 3	15	14	81%	01:09.24	00:37.20	00:44.81	00:58.14
Overall	44	41	79%	00:47.53	00:35.72	00:50.03	00:46.30

iGeology second (groups C and D)

C	D		Correct Answer Average time C	Correct Answer Average time D	Incorrect Answer Average time C	Incorrect Answer Average time D
12	9	58%	00:49.41	00:31.10	00:38.34	00:35.17
10	16	72%	00:26.88	00:36.23	00:27.02	01:15.73
16	11	75%	00:29.40	00:45.33	00:54.38	00:43.04
38	36	69%	00:35.23	00:37.55	00:39.91	00:51.31

mySoil first (groups C and D)

Correct answers	C	D		Correct Answer Average time C	Correct Answer Average time D	Incorrect Answer Average time C	Incorrect Answer Average time D
Question 1	11	12	64%	00:36.07	00:37.26	00:36.12	00:40.68
Question 2	14	11	69%	00:44.75	00:33.73	00:44.39	00:26.75
Question 3	4	14	50%	01:04.37	00:39.74	00:32.35	00:35.02
Overall	29	37	61%	00:48.40	00:36.91	00:37.62	00:34.15

mySoil second (groups A and B)

A	B		Correct Answer Average time A	Correct Answer Average time B	Incorrect Answer Average time A	Incorrect Answer Average time B
14	13	75%	00:42.89	00:39.24	00:53.88	00:27.18
11	15	72%	00:43.25	00:40.54	00:41.15	00:43.65
12	7	53%	00:33.29	00:41.42	01:10.13	00:47.54
37	35	67%	00:39.81	00:40.40	00:55.05	00:39.46

Age of participants

Cartogram

Correct answers to Question 1

	Number of participants	A	B	C	D	TOTAL		Average time A	Average time B	Average time C	Average time D	SUM Average
25-34	14	3	4	2	3	12	86%	00:30.43	00:43.52	00:26.95	00:36.81	00:34.43
35-44	22	5	3	5	2	15	68%	00:50.60	00:54.27	00:50.85	00:38.58	00:48.57
45-54	26	5	5	1	6	17	65%	00:48.52	00:39.39	00:17.12	00:39.63	00:36.16
55-64	10	1	3	4	1	9	90%	00:26.09	00:49.76	01:06.90	00:21.77	00:41.13

Correct answers to Question 2

	Number of participants	A	B	C	D	TOTAL		Average time A	Average time B	Average time C	Average time D	SUM Average
25-34	14	3	4	1	3	11	79%	00:15.87	00:28.75	00:14.76	00:24.90	00:21.07
35-44	22	4	3	5	1	13	59%	01:02.50	00:19.46	00:32.25	00:37.61	00:37.96
45-54	26	3	3	2	5	13	50%	00:51.05	00:23.02	00:19.78	00:39.93	00:33.45
55-64	10	1	2	2	2	7	70%	00:24.92	00:25.24	00:26.61	00:29.56	00:26.58

Correct answers to Question 3

	Number of participants	A	B	C	D	TOTAL		Average time A	Average time B	Average time C	Average time D	SUM Average
25-34	14	4	3	2	2	11	79%	00:30.79	00:30.19	00:20.66	00:30.18	00:27.96
35-44	22	3	3	8	2	16	73%	00:49.28	00:41.24	00:28.94	01:03.11	00:45.64
45-54	26	4	6	3	9	22	85%	00:27.28	00:41.46	00:29.15	00:41.67	00:34.89
55-64	10	1	2	3	2	8	80%	00:19.39	00:28.87	00:36.72	00:28.90	00:28.47

Overall correct answers

	Number of participants	A	B	C	D	TOTAL		Average time A	Average time B	Average time C	Average time D	SUM Average
25-34	14	10	11	5	8	34	81%	00:25.70	00:34.15	00:20.79	00:30.63	00:27.82
35-44	22	12	9	18	5	44	67%	00:54.13	00:38.32	00:37.35	00:46.43	00:44.06
45-54	26	12	14	6	20	52	67%	00:42.28	00:34.62	00:22.02	00:40.41	00:34.83
55-64	10	3	7	9	5	24	80%	00:23.47	00:34.63	00:43.41	00:26.75	00:32.06
										Grand Average		00:34.69

Age of participants

Choropleth

Correct answers to Question 1

	Number of participants	A	B	C	D	TOTAL		Average time A	Average time B	Average time C	Average time D	SUM Average
25-34	14	3	3	2	3	11	79%	00:23.25	00:28.96	00:37.96	00:23.40	00:28.39
35-44	22	5	3	4	1	13	59%	00:30.82	00:44.46	00:23.30	00:20.82	00:29.85
45-54	26	6	4	3	5	18	69%	00:28.37	00:30.85	00:44.53	00:39.84	00:35.90
55-64	10	1	3	2	0	6	60%	00:27.49	00:53.74	00:47.02	—	00:42.75

Correct answers to Question 2

	Number of participants	A	B	C	D	TOTAL		Average time A	Average time B	Average time C	Average time D	SUM Average
25-34	14	4	3	2	3	12	86%	00:34.99	00:31.42	00:31.95	00:36.80	00:33.79
35-44	22	2	5	7	3	17	77%	00:33.78	00:45.60	00:49.69	00:44.18	00:43.31
45-54	26	7	5	2	9	23	88%	00:57.19	00:35.10	00:26.81	00:35.10	00:38.55
55-64	10	1	2	3	1	7	70%	00:25.58	00:55.16	00:53.72	00:20.83	00:38.82

Correct answers to Question 3

	Number of participants	A	B	C	D	TOTAL		Average time A	Average time B	Average time C	Average time D	SUM Average
25-34	14	3	1	0	2	6	43%	01:02.52	00:25.20	—	00:34.38	00:40.70
35-44	22	6	2	3	1	12	55%	01:25.10	00:27.87	00:47.56	00:55.76	00:54.07
45-54	26	5	1	0	7	13	50%	01:01.49	00:25.58	—	00:44.44	00:43.83
55-64	10	1	3	1	1	6	60%	00:32.97	01:01.14	00:39.42	01:03.04	00:49.14

Overall correct answers

	Number of participants	A	B	C	D	TOTAL		Average time A	Average time B	Average time C	Average time D	SUM Average
25-34	14	10	7	4	8	29	69%	00:40.25	00:28.53	00:34.96	00:31.53	00:33.82
35-44	22	13	10	14	5	42	64%	00:49.90	00:39.31	00:40.18	00:40.25	00:42.41
45-54	26	18	10	5	21	54	69%	00:49.02	00:30.51	00:35.67	00:39.79	00:38.75
55-64	10	3	8	6	2	19	63%	00:28.68	00:56.68	00:46.72	00:41.93	00:43.50
										Grand Average		00:39.62

The full test results can be sourced through the following Digital Object Identifier (DOI):

10.5285/3be3b51f-f8ad-45a8-86b0-8a239de11994

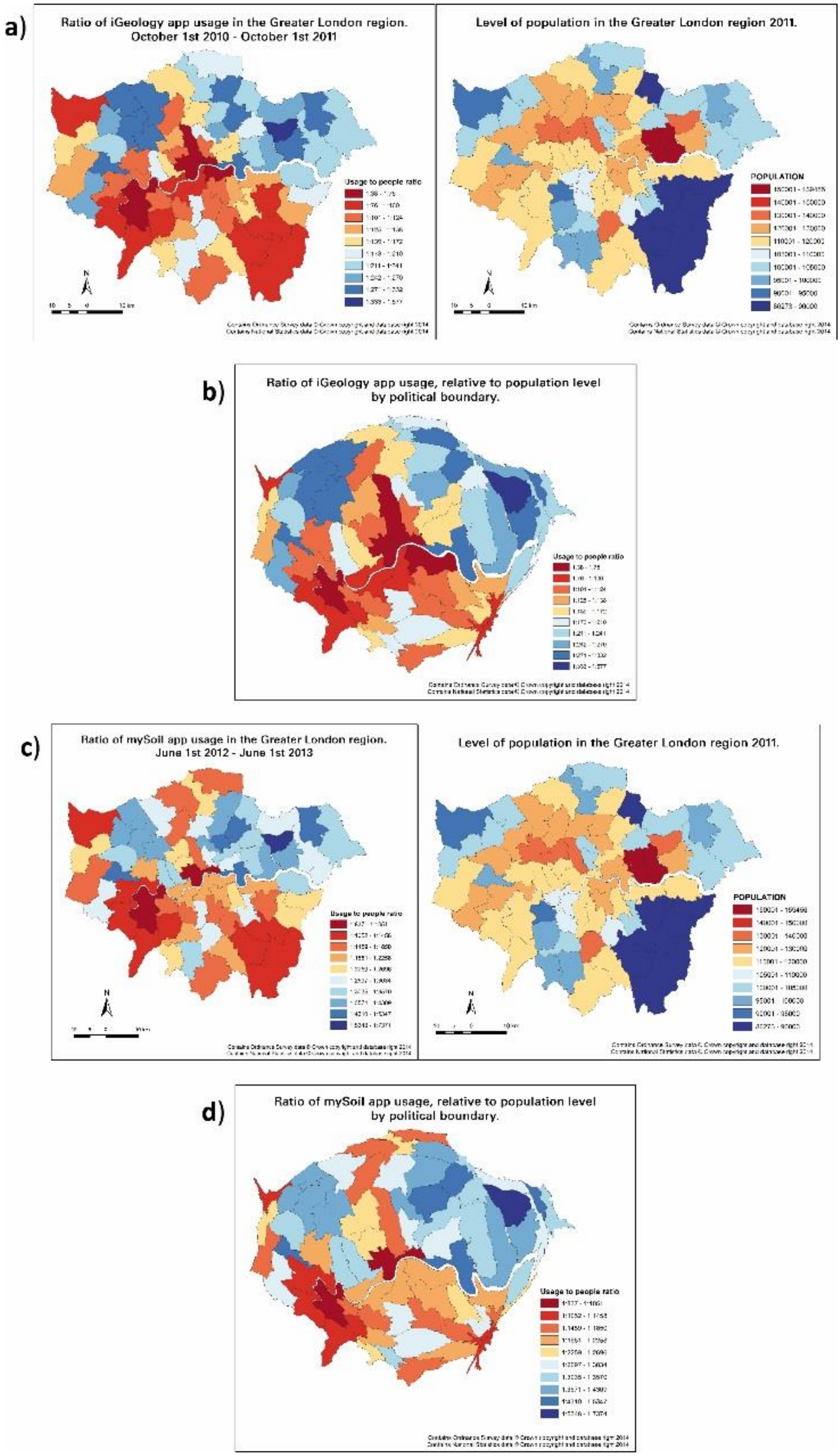


Figure 12: Maps used in the test: a) Choropleth maps using iGeology app usage data; b) A cartogram using iGeology app usage data; c) Choropleth maps using mySoil app usage data; d) A cartogram using mySoil app usage data.

7 References

British Geological Survey holds most of the references listed below, and copies may be obtained via the library service subject to copyright legislation (contact libuser@bgs.ac.uk for details). The library catalogue is available at: https://envirolib.apps.nerc.ac.uk/olibcgi.

Bortins, I., & Demers, S. (2002, 31/07/2002). Cartogram Types. Retrieved from http://www.ncgia.ucsb.edu/projects/Cartogram_Central/types.html

Buja, A., Cook, D., & Swayne, D. F. (1996). Interactive High-Dimensional Data Visualization. *Journal of Computational and Graphical Statistics, 5*(1), 78-99. doi:10.1080/10618600.1996.10474696

Crampton, J. W. (2011). *"2 Rethinking maps and identity." in Rethinking maps: new frontiers in cartographic theory* (M. Dodge, R. Kitchin, & C. Perkins Eds.): Routledge.

Cuff, D. J., & Mattson, M. T. (1982). Thematic maps: their design and production.

Cukier, K., & Mayer-Schoenberger, V. (2013). Rise of Big Data: How it's Changing the Way We Think about the World, The. *Foreign Aff., 92*, 28.

Dent, B. D. (1975). Communication aspects of value-by-area cartograms. *The American Cartographer, 2*(2), 154-168.

Dent, B. D. (1990). *Cartography: Thematic map design* (J. L. Hahn Ed. 2 ed.). Iowa, USA: Wm. C. Brown Publishers.

Dorling, D. (1996). *Area cartograms: their use and creation.* Paper presented at the Concepts and techniques in modern geography.

Dorling, D. (2015, 10th September). [(pers. comm)].

Dykes, J., & Unwin, D. (1998). Maps of the census: a rough guide. *Case Studies of Visualization in the Social Sciences, Technical Report 43 Volume, 43*.

Ferreira, J. (2014). *Infographics: An introduction*. Retrieved from Coventry University:

Galton, F. (1881). *On the construction of isochronic passage-charts.* Paper presented at the Proceedings of the Royal Geographical Society and Monthly Record of Geography.

Gastner, M. T., Shalizi, C. R., & Newman, M. E. (2005). Maps and cartograms of the 2004 US presidential election results. *Advances in Complex Systems, 8*(01), 117-123.

Goodchild, M. F. (2010). Geographic Information Systems. In B. a. Gomez & J. P. Jones III (Eds.), *Research Methods in Geography: A Critical Introduction* (pp. 376-391): Blackwell Publishing Ltd.

Griffin, T. (1983). Recognition of areal units on topological cartograms. *The American Cartographer, 10*(1), 17-29.

Guardian, T. (2014). The dimensions that shape London - mapped | UK news | The Guardian. *London: Show and Tell.* Retrieved from http://www.theguardian.com/uk-news/datablog/gallery/2014/may/11/social-dimensions-shape-london-mapped

Hennig, B. (2015). Cartograms: Human Worlds. *Geographical, 87*(5).

Hennig, B. D. (2011). *Rediscovering the World.* University of Sheffield.

Holbrook, H. W., & Cartwright, C. E. (2016). *Cartogram and Choropleth communicative effectiveness - participant test results 2015* [Dataset].

Inoue, R. (2011). GENERALIZED APPROACH TO CONSTRUCTION OF SIMPLE-SHAPED NON-CONTIGUOUS AREA CARTOGRAMS.

Inoue, R. (2011). A new construction method for circle cartograms. *Cartography and Geographic Information Science, 38*(2), 146-152.

Kardos, J., Benwell, G., & Moore, A. (2005). The visualisation of uncertainty for spatially referenced census data using hierarchical tessellations. *Transactions in GIS, 9*(1), 19-34.

Kaspar, S., Fabrikant, S., & Freckmann, P. (2011). *Empirical study of cartograms.* Paper presented at the 25th international cartographic conference, Paris.

Keserica, H., Sučić, S., & Mihajlović, Ž. (2009). *Displaying large amounts of spatial data in GIS.* Paper presented at the 32nd intenational convention on information and communication technology, electronics and microelectronics-MIPRO.

Kocmoud, C. J. (1997). *Constructing continuous cartograms: a constraint-based approach.* Texas A&M University.

Nöllenburg, M. (2007). Geographic visualization *Human-Centered Visualization Environments* (pp. 257-294): Springer.

Nusrat, S., Alam, M., & Kobourov, S. G. (2015). Evaluating cartogram effectiveness. *arXiv preprint arXiv:1504.02218*.

Nusrat, S., & Kobourov, S. (2015). Visualizing cartograms: Goals and task taxonomy. *arXiv preprint arXiv:1502.07792*.

Openshaw, S., & Alvanides, S. (2001). Designing zoning systems for representation of socio-economic data. *Time and Motion of Socio-Economic Units, Taylor and Francis, London*.

Raisz, E. (1934). The rectangular statistical cartogram. *Geographical Review*, 292-296.

Robinson, A. H., & Petchenik, B. B. (1976). nature of maps.

Shimizu, E., & Inoue, R. (2009). A new algorithm for distance cartogram construction. *International Journal of Geographical Information Science, 23*(11), 1453-1470.

Speckmann, B. (2006). Algorithms for cartograms and other specialized maps. *Geo-information and computational geometry*, 26.
Spiegelhalter, D., Pearson, M., & Short, I. (2011). Visualizing uncertainty about the future. *science, 333*(6048), 1393-1400.
Sun, H., & Li, Z. (2010). Effectiveness of cartogram for the representation of spatial data. *The Cartographic Journal*, *47*(1), 12-21.
Tobler, W. (2004). Thirty five years of computer cartograms. *ANNALS of the Association of American Geographers, 94*(1), 58-73.
van Kreveld, M., & Speckmann, B. (2004). On rectangular cartograms *Algorithms–ESA 2004* (pp. 724-735): Springer.
Weber Reuschel, A.-K., Piatti, B., & Hurni, L. (2014). Data-driven Expansion of Dense Regions-A Cartographic Approach in Literary Geography. *The Cartographic Journal, 51*(2), 123-140.
Wolf, E. B. (2005). *Creating contiguous cartograms in ArcGIS 9.* Paper presented at the Proceedings of 2005 ESRI International User Conference.