Radi(c)al Departures: Comparing Conventional Octolinear Versus Concentric Circles Schematic Maps for the Berlin U-Bahn/S-Bahn Networks Using Objective and Subjective Measures of Effectiveness

Maxwell J. Roberts,

Department of Psychology, University of Essex, UK

Elizabeth J. Newton, & Maria Canals

Department of Psychology, London South Bank University, UK

Pre-print manuscript in press:
Information Design Journal

30/11/2015

Correspondence to the first author at: Department of Psychology, University of Essex, Wivenhoe Park, Colchester, Essex, CO4 3SQ, UK, email mjr@essex.ac.uk.

Radi(c)al Departures: Comparing Conventional Octolinear Versus Concentric Circles Schematic Maps for the Berlin U-Bahn/S-Bahn Networks Using Objective and Subjective Measures of Effectiveness

Abstract

An experiment is reported in which two designs of Berlin U-/S-Bahn network maps were compared for usability. One was conventional, based on standard schematic design rules used worldwide: Straight lines with tightly radiused corners, and only horizontal, vertical, or 45° diagonal angles permitted. The other was a novel concept, based on concentric circles and spokes radiating from a central point. The former has the benefit of simple line trajectories, the latter potentially has the benefit of a coherent overall appearance. The experiment investigated both an objective performance measure (time required to plan complex journeys) and a variety of subjective measures (choice between maps, ratings of statements associated with usability, direct ratings of usability). All subjects planned journeys using both designs. Overall, performance was worse for the concentric circles map, and it received poor ratings. However, in line with previous research, objective and subjective measures were dissociated. For example, many subjects expressed a preference for the design that was not the best for them in terms of objective performance.

Running Head

Octolinear vs. Concentric Circles Schematic Maps

Keywords

Schematic maps, Metro maps, Concentric Circles maps,
Octolinear maps, Journey planning, Usability study,

Radi(c)al Departures: Comparing Conventional Octolinear Versus Concentric Circles Schematic Maps for the Berlin U-Bahn/S-Bahn Networks Using Objective and Subjective Measures of Effectiveness

1. Introduction

Urban rail networks are increasing in complexity worldwide, owing both to lines and extensions being added within existing networks (particularly in Asia), and also to the integration of different transport modalities within cities, made possible by common fare arrangements. Ovenden (2003, 2007, 2015) provides an excellent time-lapsed overview of this phenomenon. With this increasing complexity, the user is faced with ever more options for journey planning: Identifying appropriate departure and arrival points, and determining the most efficient route in order to travel from one to the other. One solution to this problem, widely adopted, is the creation of online computerized journey-planners. These have the advantage of being able to take account of real-time data, but have the disadvantage that the user is rarely presented with an overview of the network, from which its structure can be learnt. However, more traditional aids such as static network maps are still produced, and Ovenden (2015) continues to compile a set of these for every major rail-based urban transit undertaking.

A popular information design technique for the creation of network maps is to use a schematized representation. For this, complex line trajectories are simplified, often at the expense of topographical precision, and surface details minimized. Again, Ovenden (2003, 2007, 2015) shows clear trends. Small and/or recently constructed systems tend to provide topographical maps. For these simple, compact networks, there is little pressure on designers to adopt schematization techniques, and the immaturity of the networks may mean that users need assistance in embedding the locations of the facilities within the wider context of the topography of the city. Larger and/or more mature networks, such as London, Berlin and Madrid, tend to have highly schematized maps. For these, showing outlying suburban destinations may require scale distortion (resulting in compression) and the high station density at central focal points may require enlargement for clarity. Together, these balance the design in terms of data density. In tandem with this, greater user-knowledge of the structure of a mature network may reduce the need to position facilities precisely in line with city topography, enabling an opportunity to simplify line trajectories.

Defining whether a network map can be categorized as schematic versus topographical is not always easy (Dow, 2005; Roberts, 2005). The London Underground network diagram, originated by Henry Beck and first issued in 1933, is the most historically and internationally well-known and prototypical example (Garland, 1994). For this, street details are absent and the River Thames is the only landmark shown. Straight lines with tightly radiused corners are used to show services, and only horizontal, vertical, and 45º diagonal angles are used. The simplification of line trajectories is associated with topographical distortion, although considerably less for Beck's earlier work than his later designs (Roberts 2005, 2012). Looking at other network diagrams worldwide, these schematization techniques are in evidence, although other surface details such as parks and major streets may be included, and topography may be distorted to a greater or lesser extent. Very few schematic maps depart from the angles preferred by Beck, and this method of design is also known as *octolinearity* (or *octilinearity,* Nöllenburg & Wolff, 2011; Wolff, 2007): When creating such a map, at any given point, the designer has only *eight* different directions to choose between.

1.1 Prescriptions for Designing Effective Schematic Maps

Octolinearity has become something of a design gold standard internationally, such that there is widespread belief amongst graphic designers, researchers, commentators, and users that applying this will result in the best design possible no matter what the structure of the network (e.g., Ovenden, 2005, p. 39). For example, until recently, this dictated the objectives of researchers attempting to automate schematic map design, so that Nöllenburg & Wolff (2011) described octolinearity as a *Hard Constraint* (i.e., it should never be broken) and suggest that "the main benefit of octilinear layouts is that they potentially consume less space and use fewer bends while still having a tidy and schematic appearance" (p. 626) and that "we believe that octilinearity, which is strictly followed by most real metro maps, is an essential ingredient for tidy and easy-to-read metro map layouts" (p. 627). Choice of angles is the most fundamental determinant of the structure of a schematic map, and so the plausibility of this stance should be the first consideration in any analysis of design effectiveness.

Roberts *et al*. (2013) criticize the *octolinearity as a gold standard conjecture* from a number of different viewpoints. First, they note that there is virtually nothing in the psychological literature to predict or corroborate this. Second, belief in this conjecture discourages consideration of the possibility that networks with different structures may require different design solutions, and also deflects from analysing maps from the perspective of the quality of implementation within the design rules adopted. Third, in a series of comparisons between the official octolinear Paris Metro map and a novel curvilinear design, Roberts *et al*. (2013) showed that the curvilinear version was always associated with considerably improved journey planning times (up to 50% faster, see Figure 1). This refutes the *strong* version of the gold standard conjecture (octolinearity will *always* result in the most effective design) although weaker versions are left intact (e.g., octolinearity will *usually* result in the most effective design, but not in instances where this is incompatible with network structure).

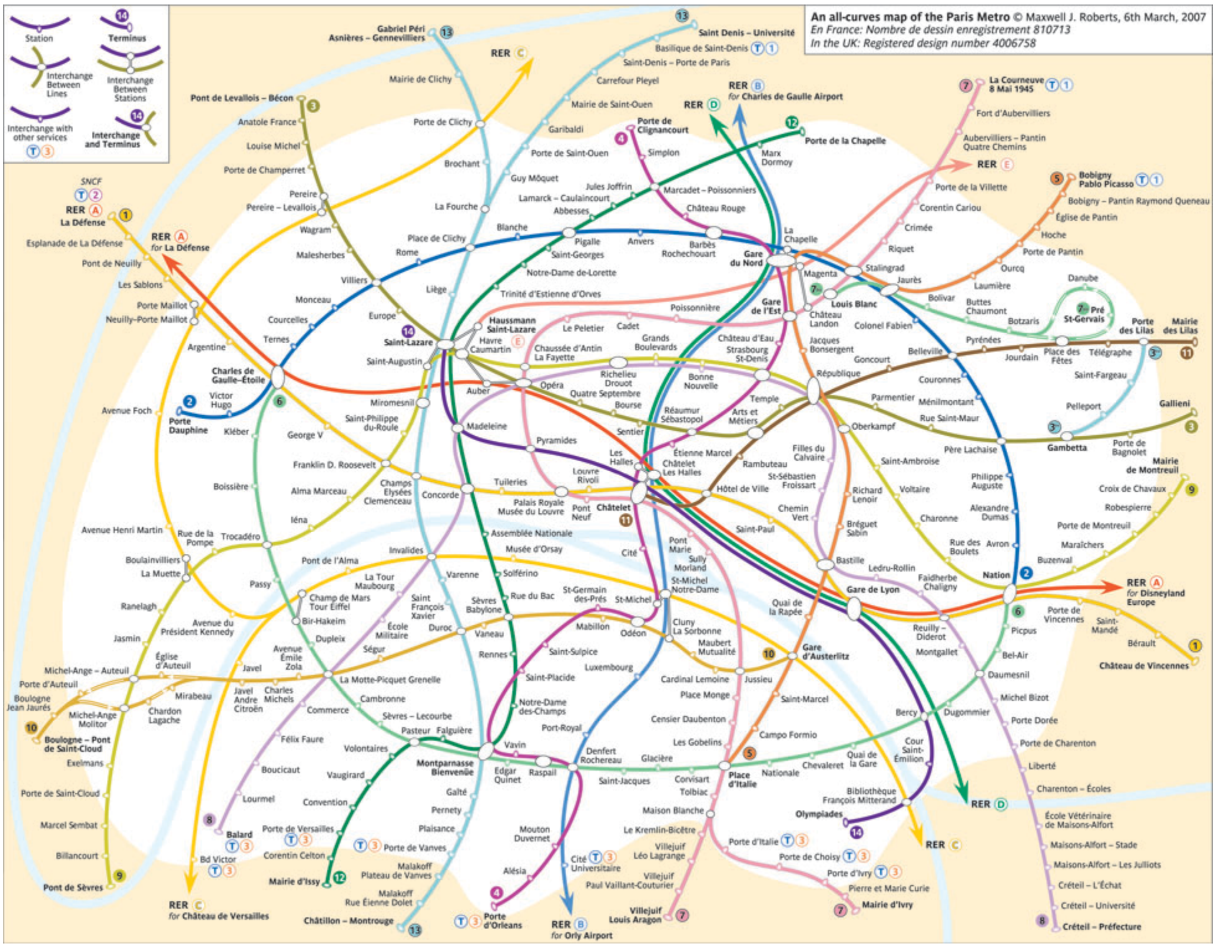

Figure 1. The curvilinear map used in the study by Roberts *et al.* (2013).

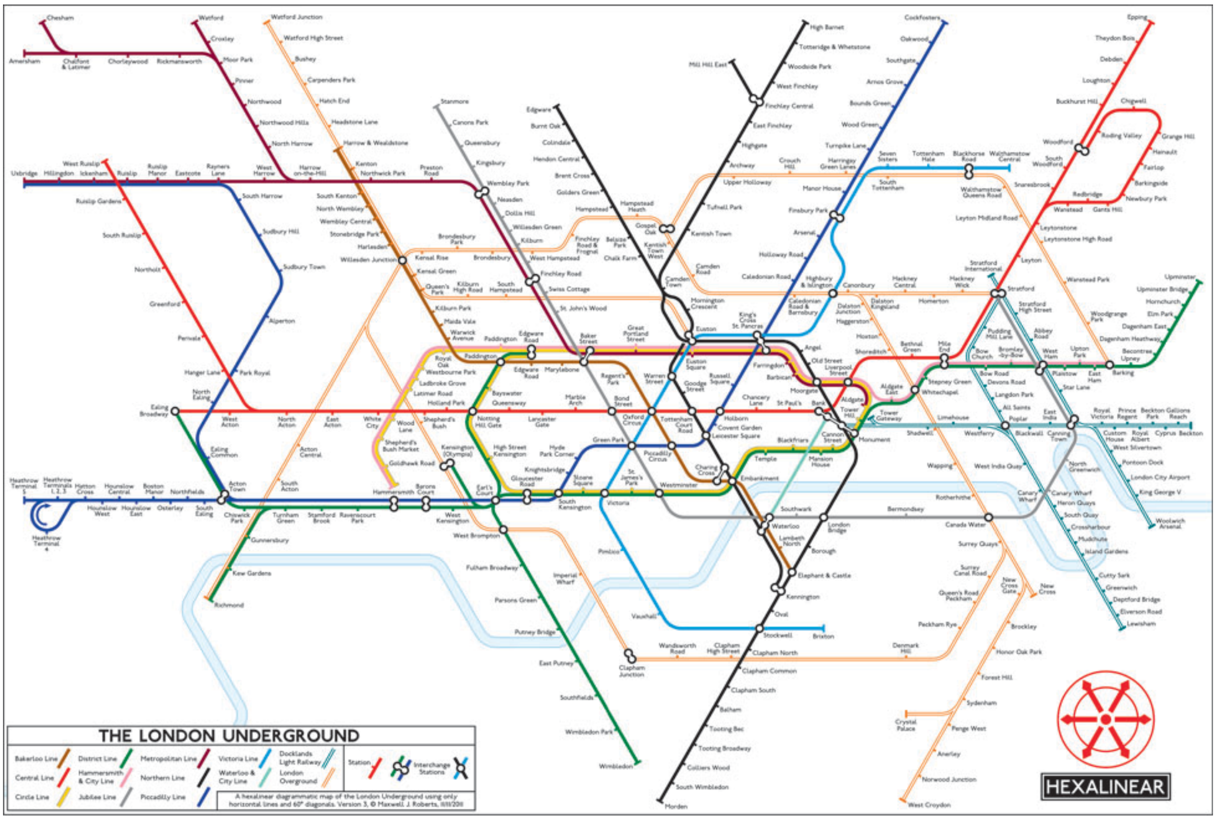

Figure 2. A hexalinear map of the London Underground. These angles are particularly compatible with the structure of the network inside the Circle Line, giving simpler line trajectories in this region than is possible with an octolinear design.

Roberts (2012) notes that octolinearity is just one of a sequence of angle-sets that could be used for schematic maps. In other words, the *level of linearity* can be varied. Hence, a design might use just two perpendicular angles (tetralinear), three angles at 60° to each other (hexalinear, see Figure 2), four angles (octolinear), five angles (decalinear) and so on. A systematic and exhaustive exploration of these is suggested for any city whose network is to be mapped, and Roberts (2012) implements this for Berlin and London. The aim of this procedure is for the designer to specify the objectives for a design in advance, and then identify the level of linearity that best-enables these to be met. However, precise guidance on design objectives is remarkably rare (Roberts, 2014b). Nöllenburg (2014) gives a comprehensive set (e.g., keep line trajectories as straight as possible, space stations evenly, station labels should not occlude lines, and relative positions of stations should be preserved). Ovenden (2008) provides a set that is broadly compatible, but includes more subtle prescriptions (e.g., do not bend a line twice between a pair of stations, keep station labels horizontal). Many of these seem reasonable, and can shown to be compatible with theories of human cognition (see later). However, some criteria are somewhat subtle, and are more likely to have implications for aesthetic preference than measurable effects on journey planning. In any case, few of the individual prescriptions have direct empirical support, although Newton & Roberts (2009) report a study that shows that more complicated London Underground maps (in terms of bends in line trajectories) are associated with longer journey planning times. One problem with empirical studies to date on schematic map usability is that they have focused on comparing very differently conceived designs (e.g., Bartram, 1980; Bronzaft & Dobrow, 1984; Bronzaft, Dobrow, & O'Hanlon, 1976; Roberts *et al*., 2013; Rosenholtz, 2011; for a review, see Roberts, 2014a) rather than systematic investigations of variations in design rules and objectives.

1.2 A Framework for Effective Schematic Map Design

Roberts (2012, 2014b) attempts to organize the various criteria for effective design into a broad framework of five categories. The prime consideration of these is the configurational issues of design.

Simplicity. Individual lines should have simple, easy-to-follow trajectories with minimal changes of direction.

Coherence. Collectively, lines should relate to each other to give the overall diagram clear pattern and good shape, for example by maximizing parallel lines.

Balance. Ideally, there should be an even density of stations across a map, or at least gentle density gradients from a dense centre to sparse suburbs.

Harmony. Roberts (2012, 2014b) suggests a placeholder category for aspects that are likely to influence the aesthetics of a design (albeit with substantial individual differences; Palmer, Schloss, & Sammartino, 2013) but are unlikely to have any measurable impact on usability.

Topographicity. A schematic with poor topographicity is one in which the topographical distortion is sufficiently extreme that it adversely affects user-confidence in the design, or worse, leads to the planning of inefficient or inappropriate journeys (Guo, 2011).

From a psychological perspective the need for, *simplicity, coherence,* and *balance* are easy to justify. The perceptual qualities of maps has been investigated (e.g., Rosenholtz, 2011) and the balance of a design would be expected to impact on its attention capture. For example, a map with a very diffuse, open centre and dense, packed suburbs around the periphery, would lack a clear focus, with attention incorrectly directed towards high station densities at the edges. Aspects of higher cognition are also likely to play a role. For example, the difficulty of intelligence test items is related to: (1) the *simplicity* of the items in terms of the numbers of elements that they contain; (2) the *visual complexit*y of the elements; and (3) the *coherence* of the configurations of elements in terms of how they are organized holistically (Meo, Roberts, & Marucci, 2007; Primi, 2002; Vodegel Matzen, Van der Molen, & Dudink, 1994; Roberts, Livermore, Welfare, & Theadom, 2000). These factors affect performance by raising the *cognitive load* of solving a problem: The more information necessary to process in order to perform, the harder the task becomes (Carpenter, Just, & Shell, 1990). These factors also affect performance because it is harder to undertake a task if its key elements, and their relatedness, are difficult to identify (Meo et. al., 2007; Roberts *et al*., 2000).

Overall, Roberts *et al*. (2013) argue that a poorly-designed schematic map buries its elements (the lines) in visual complexity, making them hard to identify, and analogous to a difficult intelligence test. A simple, coherent, balanced schematic map will have high *structural salience,* revealing the elements and their relatedness, and making the underlying network structure easier to identify, facilitating both journey planning and learning, so that a virtuous circle is set up, with performance getting better as more is learnt. For such a design, we would therefore expect fast journey planning, few errors, better remembered plans, and more easily reconstructed plans in the event of a failure to remember. In comparison, a poorly designed schematic will not have these benefits, and may even have little to offer compared with a topographical map, other than the simplification entailed in removing street details and most other landmarks.

1.3 Beyond Octolinearity

A key feature of the framework put forward by Roberts (2012, 2014b) is that it is neutral with respect to design rules. Hence, these should be selected such that they are compatible with the network structure, enabling the framework criteria to be best-optimized. Provided this is achieved, then the level of linearity of the design does not matter, and even more radical departures might be appropriate. The motivation behind the curvilinear Paris Metro schematic investigated by Roberts *et al*. (2013) was that this network is uniquely disorganized and interconnected, and the high station density across the system limits the scope for simple octolinear line trajectories unless there is excessive topographical distortion. It was argued that under these circumstances, any linear map would be hard-pressed to achieve simplicity and coherence, and therefore a curvilinear design, using only Bézier curves, should be treated as a serious alternative, replacing continual disruptive zigzags with gentle curves. Bézier complexity can be reduced by avoiding variations in curvature, sharp corners, control points, and points of inflexion. The design easily outperformed the official octolinear version.

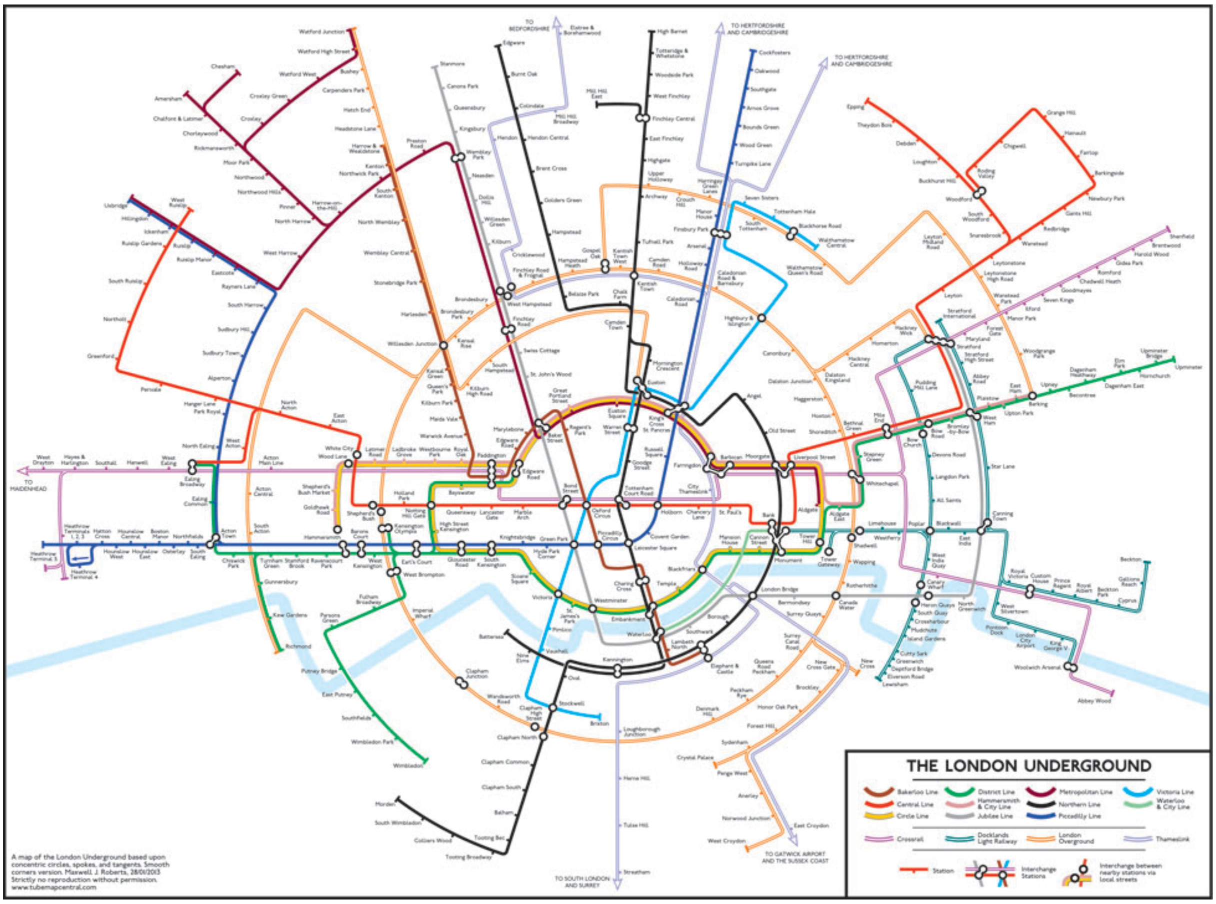

Figure 3. A Concentric Circles map of the London Underground, which generated a massive response in the British media.

Another radical departure from traditional design approaches is to base the map solely on concentric circles and spokes radiating from a central point. The result of such an exercise, carefully executed, can be a visually powerful design. This technique was investigated for the London Underground network, by the first author, as an intellectual exercise (See Figure 3). The response to this via the Internet was unexpectedly positive. The design was featured on numerous blog and media web pages, and concentric circles maps for New York and Berlin generated similar positive responses (e.g., Creative Review, 2013; Die Welt, 2013; Die Zeit, 2013; Gothamist, 2013; ITV News, 2013; Metro, 2013). Commentators frequently express an opinion that such designs are clearer than official alternatives, and the author has received a request to reproduce the Berlin design in a city travel guide. Researchers into automated schematic map design have now implemented algorithms for creating concentric circles maps (Fink, Lechner, & Wolff, 2014). From the point of view of the usability framework earlier, such designs have poor simplicity: The restricted rules mean that complex line trajectories are necessary. Conversely, such designs may have high coherence, forcing city networks into unprecedented levels of organization. This analysis forms the basis of the study reported here, comparing usability and preference between a concentric circles design and one using traditional octolinearity.

Concentric circles maps appear visually powerful, generating strong positive reactions amongst many users. However, there are good reasons to treat these responses with caution. A recurring feature of research into schematic map design is that subjective ratings of usability have little or no correlation with objective measures. People can rate maps highly that they previously found difficult to use, and *vice-versa*. Hence, in the Paris Metro study earlier, Roberts *et al*. (2013) found that despite the superiority of the curvilinear map over the octolinear version in terms of objective measures, there were no differences between them for questionnaire ratings or preferences. Roberts (2014b), in a preliminary analysis of a rating study of alternative London Underground maps, found that users were sensitive to differences in map simplicity within a set of design rules. For example, octolinear maps with simple line trajectories were rated as more usable and more attractive than octolinear maps with complex trajectories. However, ratings of matched maps with different design rules tended to be biased towards octolinear designs, considerably more so than would be expected on the basis of usability studies. Again, these findings are consistent with psychological literature. People tend to be poor observers of their own performance and have little insight into their own cognitive processes (e.g., Chabris & Simons, 2010; Kruger & Dunning, 1999). Without explicit feedback, a user would simply be unaware of his or her own performance in terms of the time required to plan each journey. People are also susceptible to biases generated by expectations and prejudices. For example, the *mere exposure* effect is well-documented in psychology (e.g., Bornstein, 1989). Repeated exposure to, and increasing familiarity with stimuli, results in more positive ratings compared with less familiar material. Furthermore, an important finding in the expertise literature is that novices in any domain tend to evaluate items according to superficial surface properties (e.g., Chi, Feltovich, & Glaser. 1981). Hence an octolinear map might be over-favourably evaluated by a person who is familiar with such designs, but not an expert at usability issues.

Taken together, the various competing factors mean that it is virtually impossible to predict the usability of concentric circles maps. Their poor simplicity of line trajectories may be compensated for by high overall coherence. Subjective responses have not been gathered formally, and so the positive reactions are even less trustworthy than those ascertained by questionnaires under laboratory conditions. Even so, these responses have been stronger than for any other experimental design released on the Internet and, presumably, have been expressed despite people’s general expectations that schematic map design should be octolinear for maximum usability. However, another important consideration is that people tend to have aesthetic preference for curved objects as opposed to objects with sharp contours (e.g., Leder, Tinio, & Bar, 2011; Silvia & Barona, 2009). Hence, positive responses could reflect aesthetic evaluations of such designs, rather than assessments of usability.

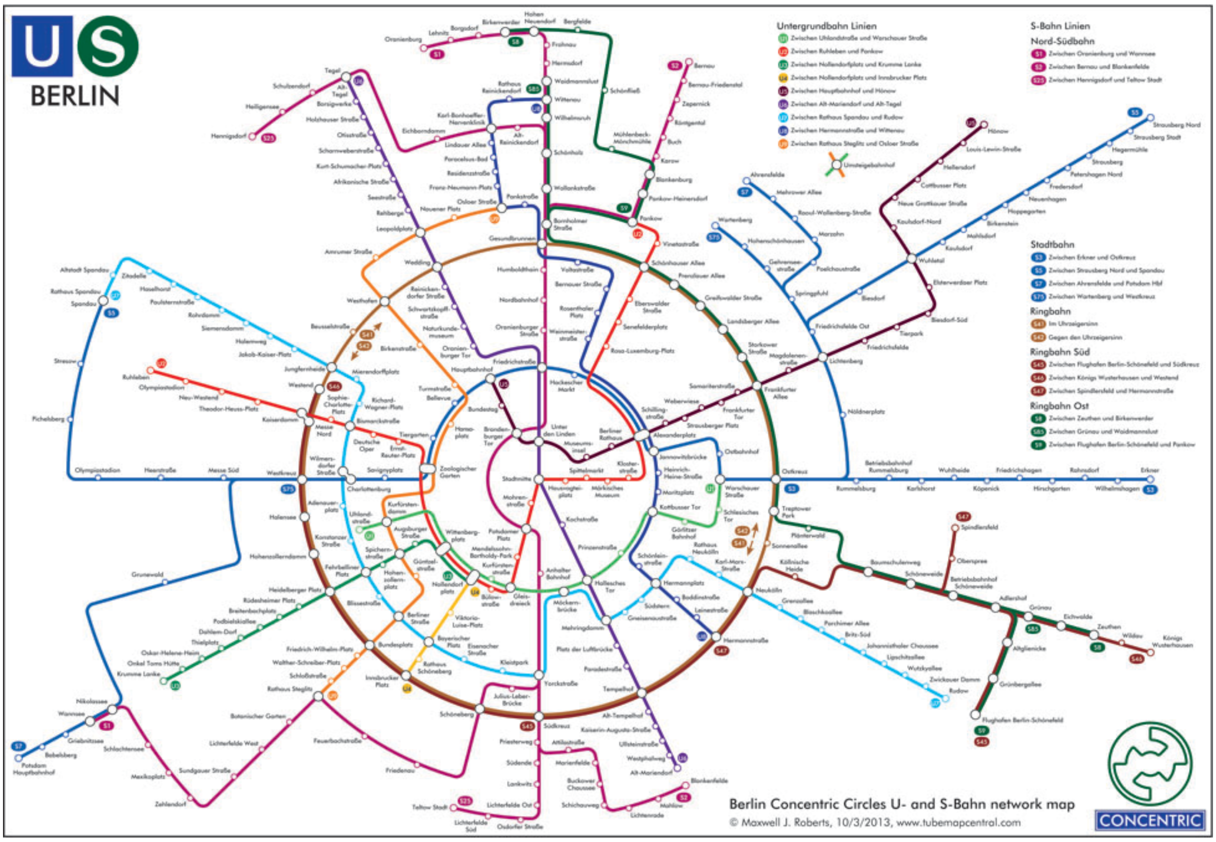

Figure 4. The concentric circles map used in this study.

1.4 The Study Reported Here

This study compares two schematic maps depicting the Berlin U-Bahn and S-Bahn network: one concentric circles-based (Figure 4), the other traditional octolinear (Figure 5). Both were designed manually by the first author using computer vector graphics packages, and matched for colour-coding, font size, and surface area.[1] Berlin was chosen for a number of reasons. It has a well-known circular line (S41/S42) and hence a concentric circles map is conceptually appropriate. For this city, the method of design presented few challenges; the structure of the network is not incompatible. Alternative cities were ruled out for various reasons. For example, London (the structure of the network dictates breaking the requirement that all straight line-spokes must radiate from the centre) Paris (a very unbalanced design results) New York (there is no orbital line) and Moscow (the network is highly interconnected so that few indirect journeys require more than one interchange). The structure of the Berlin network means that many journeys can be devised that are complex in the sense that two interchanges will be necessary to complete them. This maximizes the likelihood of finding usability differences between designs. The octolinear map tested was not the official version, which is poorly optimized for simplicity of line trajectories. Both experimental designs omitted the long-distance rail services that are present on the official version, and the line trajectories of the octolinear design were optimized for simplicity. Participants were resident in London, and thus would be expected to be particularly familiar with octolinear schematics, and unfamiliar with the Berlin network.

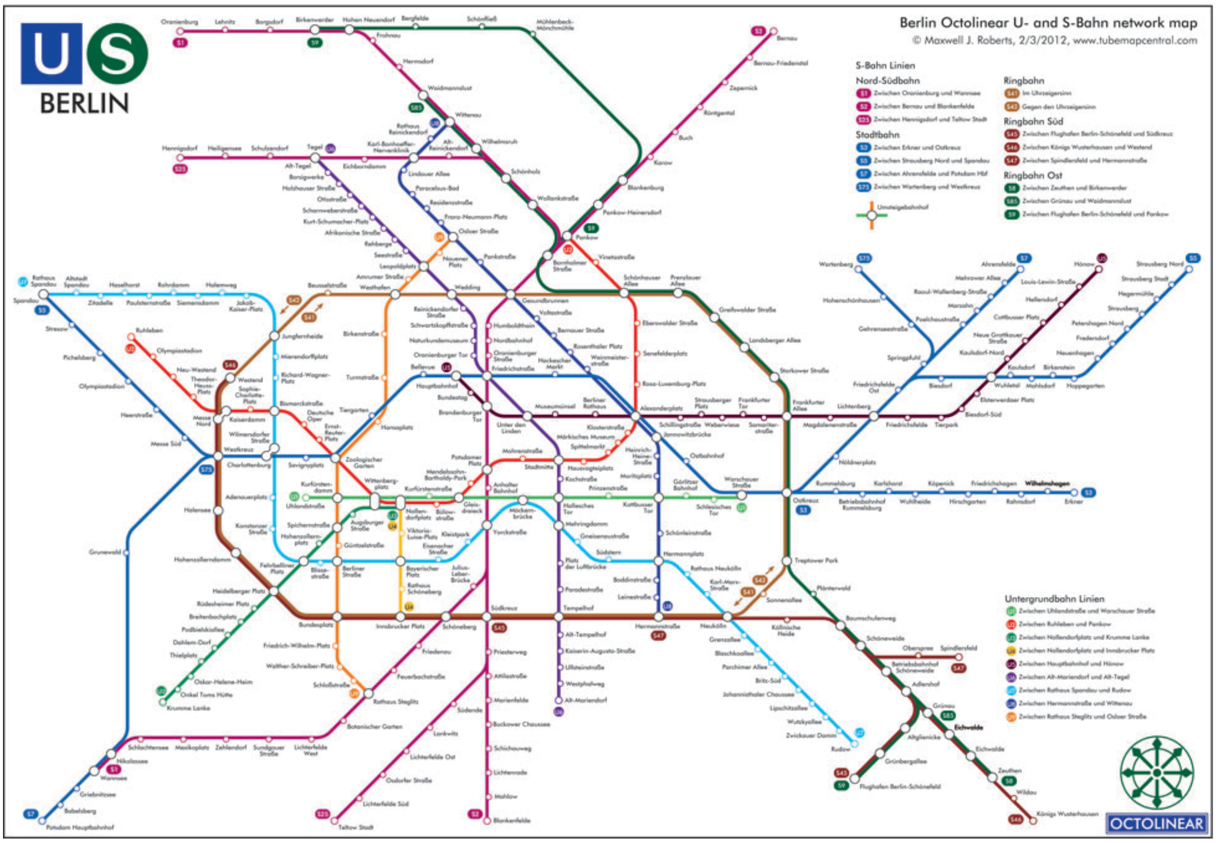

Figure 5. The octolinear map used in this study.

Subjects were asked to plan various complicated journeys (i.e., minimum two interchanges required) using both designs, one after the other. Each person therefore had experience with both maps, enabling the best-performing design for each individual to be identified. This is the first time that such a within-subjects design has been reported for a schematic map usability study -- Roberts *et al*. (2013) used a between-subjects design in which each person planned journeys using just one map, which limits the scope for comparisons between designs for individuals. A similar methodology was otherwise used to Roberts *et al*. (2013): Usability measures included journey planning times, and planning errors (e.g., if a journey included an attempt to change between lines at a point where no interchange was shown), although it was not expected that either design would yield many errors. Journey durations were also estimated, primarily as a control variable, to ensure that, for example, a tendency for faster planning for one map was not traded off against longer, more roundabout journeys. Overall, journey planning time was the prime variable of interest. A questionnaire, similar to Roberts *et al*. (2013) was administered after the journey planning task, so that subjects were asked to judge each map according to a number of questions on different aspects of design and usability. Finally, direct usability and attractiveness ratings were solicited, along with a request to express a preference for one map over the other.

2. Method

2.1 Subjects

Forty subjects took part in this experiment, 19 males and 21 females. They were unpaid volunteers from London South Bank University with a mean age of 30 years (*SD* 7). Twenty-six of them had no previous experience with the Berlin network.

2.2 Materials

The maps for journey planning were printed to fit an A3 sheet and laminated. Three sets of journeys were assembled as follows:

Set 1
S8 -- U4 Mühlenbeck-Mönchmühle to Rathaus Schöneberg
U7 -- S5 Eisenacher Straße to Birkenstein
U1 -- S2 Görlitzerbahnhof to Mahlow
U5 -- U3 Kaulsdorf-Nord to Hohenzollernplatz

U8 -- U2 Paracelsus-Bad to Olympiastadion
U9 -- S25 Güntzelstraße to Eichborndamm

Set 2
S85 -- U9 Plänterwald to Hansaplatz
U6 -- U8 Westphalweg to Rosenthaler Platz
U3 -- S1 Thielplatz to Borgsdorf
S8 -- U7 Bergfelde to Südstern
S2 -- U2 Zepernick to Neu-Westend
S46 -- S5 Wildau to Strausberg

Set 3
U4 -- S5 Viktoria-Luise-Platz to Neuenhagen
U9 -- U5 Walther-Schreiber-Platz to Louis-Lewin-Straße
U7 -- S1 Haselhorst to Frohnau
U8 -- U3 Rathaus Reinickendorf to Onkel Toms Hütte
U1 -- S47 Kurfürstenstraße to Oberspree
S7 -- S8 Babelsberg to Schönfließ

Journeys were chosen such that they would be difficult to plan, with distant start and destination stations and many alternative options. Two interchanges were always required either to complete the journey at all, or to avoid a roundabout journey. An attempt was made to represent all lines/regions of the map equally between item sets. Each journey was shown on an individual A4 laminated sheet with the map greyed out except for the start (arrowed) and destination stations.

A 22-item questionnaire was created, based on Roberts *et al*. (2013). The main difference was that five point rating scales were used where appropriate (as opposed to seven points previously). For statement rating questions, the scale offered the options of *strongly agree/agree/neutral/disagree/strongly disagree*. For every question, a separate answer or a decision was required for each of the maps. The full set of questionnaire items was as follows.

Questions 1 to 15 were rating-scale questions as described previously. Asterisks denote questions which directly ask for opinions about aspects of design related to usability. This part of the questionnaire will be referred to subsequently as the *statement rating task*.

**1) I found journeys easy to plan using this map*
**2) Routes were difficult to discriminate (identify) using this map*
**3) Station names were easy to identify using this map*
**4) Station interchanges were difficult to negotiate using this map*
**5) Line trajectories were easy to follow using this map*
**6) I found this map disorientating to use*
**7) I would be happy to use this map to plan real-life journeys around Berlin*
8) I preferred a direct-looking route, no matter how many interchanges required
9) Some parts of the map looked complicated, and I planned journeys to avoid them
10) This map is intended for planning journeys but I think it is also geographically accurate
**11) With this map, I would rather walk or take a taxi than use the Berlin U-Bahn*
12) The best routes for me had the fewest station stops along the way
**13) I found the map visually disturbing*
**14) I found the map clean and uncluttered*
**15) I would look for another design of Berlin U-Bahn map to use at the earliest opportunity*

Questions 16 and 17 requested brief sentences, i.e., *qualitative* responses.
16) Briefly, what, if any, aspect of each map did you like the most?
17) Briefly, what, if any, aspect of each map did you like the least?

Question 18 was a forced choice item, one map or the other preferred. This will be referred to subsequently as *map choice*.
18) Of the two designs you have used, which one do you think you would prefer for everyday use?

Questions 19 and 20 gave a range of frequency options.
19) Roughly how often do you travel by rail to make a journey in a town or city?
20) Roughly how often have you visited Berlin?

For questions 21 and 22, for each map, for each question, there were three options: *easy to use/neutral/hard to use* or *attractive/neutral/unpleasant*. These will be subsequently referred to as the *usability rating* and the *attractiveness rating*.
21) For each of three Berlin maps, please summarize how you feel about it by deciding whether, overall, you think that it is easy or difficult to use
22) For each of the three Berlin maps, please summarize how you feel about it by deciding whether, overall, you think that it is attractive or unpleasant to look at

2.3 Design

All subjects planned journeys using both test maps, six journeys for one design, then six journeys for the other. Nineteen subjects planned journeys using the octolinear map first, the remaining 21 received the concentric circles map first. Primarily, this was a within-subjects design with Map Type (two levels, Octolinear versus Concentric Circles) as the independent variable. Measures of map performance included the time taken to plan a journey, planning errors, and an estimation of the duration that the planned journeys would have taken had they been implemented. Questionnaire data provided a means of assessing people's subjective ratings of map usability.

Each subject was asked to plan journeys taken from two of the three sets of items. For example, Set 1 with the octolinear map, and Set 3 with the concentric circles map. The journey set selected, along with the order of presentation of maps, was counterbalanced to the greatest extent possible given the sample size. The order in which the maps were presented was also analysed to ensure that one order combination did not result in different patterns of performance to the other. Individual journeys were also analysed to ensure that any differences in overall usability between maps were not due to idiosyncrasies of just a few individual journeys.

2.4 Procedure

Subjects were tested individually at a desk or table in quiet surroundings. They were informed that they would be asked to plan a series of Berlin journeys using the supplied maps. They were to assume that the network was fully operational and that there were no cost considerations. They were given no guidance as to journey criteria or priorities, it was simply stated that they should devise the journey that they would choose if they were actually to undertake it in real life. They were also informed that they should only change between lines at designated interchanges shown on the map.

Subjects were given an opportunity to view the first map while the initial instructions were given. One practice journey was administered, chosen at random from the unused set of items. There then followed the six test journeys presented in a random order. Each trial commenced with the experimenter placing the journey sheet indicating start and end stations above the A3 laminated map, and timing immediately commenced using a stop-watch. The participant was asked to plan the journey requested, using a dry-wipe marker. Once satisfied with the plan, a verbal announcement was made by the participant, timing stopped, and the final chosen route was transcribed onto an A4 paper map by the experimenter. Following this, the experimenter cleaned all marks from the laminated map and the next trial commenced. Once all six journeys were completed, the process was repeated with the second map, including a second practice trial chosen randomly from the unused set. When all journeys had been planned, participants completed the questionnaire with both maps in sight.

Table 1. Mean usability measures by Map Type, and aggregate questionnaire ratings, and also performance by map presentation order and map preference.

		Octolinear	Concentric Circles
Planning time (seconds per journey)	Mean	25.2	30.9
	SD	7.6	9.6
Estimated journey duration (minutes)	Mean	62.5	62.4
	SD	6.6	5.9
Aggregate statement rating task score (11 to 55, high scores better)	Mean	44.4	33.7
	SD	7.2	10.4
Planning time (seconds per journey) Octolinear map presented first (*N*=19)	Mean	25.1	29.5
	SD	6.5	7.7
Planning time (seconds per journey) Concentric Circles map presented first (*N*=21)	Mean	25.3	32.2
	SD	8.7	11.1
Planning time (seconds per journey) Octolinear map is chosen (*N*=33)	Mean	26.0	32.0
	SD	7.7	9.7
Planning time (seconds per journey) Concentric Circles map is chosen (*N*=6)	Mean	22.8	27.5
	SD	5.9	6.9
Aggregate statement rating task score Octolinear map is chosen (*N*=33)	Mean	45.4	31.2
	SD	6.8	9.5
Aggregate statement rating task score Concentric Circles map is chosen (*N*=6)	Mean	37.7	44.0
	SD	5.7	5.3

3. Results

3.1 Objective Usability Measures

For each subject, means of the planning times were calculated for the six journeys for each map. For each journey, its duration was estimated by allowing two minutes per station and ten minutes per interchange. This is comparable with the heuristics that passengers themselves use (e.g., Vertesi, 2008) and ignores the variable interchange quality within most metro networks, which is virtually impossible to communicate via maps. The value of ten minutes is a worst-case scenario estimation, should an unknown interchange prove to be particularly long, and/or a train is just missed (see Guo & Wilson, 2011). No subjects made any planning errors, and therefore these are not reported. Mean performance by map is shown in Table 1.

The effects of Map Type on the usability measures were analysed using single factor repeated measures Analysis of Variance. There was a substantial difference in the mean time taken to plan journeys between maps, with better performance for the octolinear version, and this was statistically significant, $F(1,39) = 46.1$, MSe = 14.4, $p < .01$. The negligible difference in mean estimated journey duration was not statistically significant, $F(1,39) = 0.01$, MSe = 43.9, $p > .05$. Hence, the faster planning times for the octolinear map were not associated with poorly-planned journeys of a longer estimated duration.

Table 1 also shows that, despite each subject planning twelve journeys using two different maps, there is little evidence for order effects: Experience with one design had little influence on performance at the other. For example, mean planning time for the octolinear design was around 25 seconds per journey whether the map was used first, or after experience with the concentric circles design. The lack of order effects was confirmed by performing a two-factor mixed design Analysis of Variance with Map Order as the between-subjects factor, and Map Type as the within-subjects factor. The interaction between the two factors was non-significant, $F(1,38) = 2.35$, MSe = 13.9, $p > .05$.

It is possible to investigate whether there are individual journey effects that are driving the difference between designs. For example, the concentric circles map and the octolinear map might usually be equivalently easy for planning, except for one or two journeys that are particularly difficult for the concentric circles map, which have a disproportionate effect, skewing the overall means. This possibility is unlikely because the trend is highly consistent from person to person: 34 out of 40 subjects had a faster journey planning mean for the octolinear than the concentric circles design, and this finding was derived from samples taken from eighteen different journeys. If the mean time is calculated for each individual journey for each map, 17/18 of these have faster planning times for the octolinear design compared with the concentric circles map, with a median difference of five seconds in favour of octolinear. This value is close to the difference in overall means between maps (5.7 seconds, see Table 1) indicating that performance is unlikely to have been skewed by a small number of outlier journeys that were particularly disadvantageous for the concentric circles map.

Overall, looking at objective measures of performance: In terms of the time taken to plan journeys, the octolinear design has a clear and unqualified advantage over the concentric circles map, being around 20% faster. This effect is highly consistent both looking at individual subjects and individual journeys. Also important, these findings demonstrate the viability of a within-subjects design experiment for investigating different maps: There was little or no evidence for performance at one contaminating performance at the other.

3.2 Subjective Usability Measures

Subjects were given the opportunity to express opinions on the maps in a variety of different ways via the questionnaire. The most straightforward was asking them to simply to choose which of the two maps they would prefer to use in a real setting, the *map choice*. Just 6/40 people expressed a preference for the concentric circles design. The *statement rating task* on the questionnaire permits a more sophisticated analysis. For this, an aggregate score was created using the 11 questions that are directly relevant to usability issues. The questions were designed to be bi-directional, so that agreeing with the statements in some questions indicated a positive assessment of a design, but agreeing with the statements of other questions indicated an adverse assessment. Scores on the latter type were reversed (so that a rating of 1 became 5, 2 became 4, etc.) and the scores for each person, for each map, were then totalled separately. This gave aggregate statement rating task scores, in which 55 indicated the highest possible rating of a design, 11 the worst possible, and 33 a neutral rating. Table 1 shows that the octolinear map was given a higher mean aggregate rating than the concentric circles design, and the difference was significant, $F(1,39) = 27.0$, $MSe = 85.6$, $p < .01$. Looking at individual questions, the concentric circles map was generally less adversely rated for fine details such as station name identification, interchange navigation, discriminating between lines and following individual ones, but was particularly adversely rated for more holistic measures, such as the design being disorientating and visually disturbing overall.

At the end of the experiment, subjects were asked to rate each of the designs directly for usability and attractiveness, both aspects on a simple three point scale. The responses to the *usability ratings* were scaled such that if a design was rated as easy to use by everyone, then an overall value of 100% would be attained, and if a design was rated as hard to use by everyone, then an overall value of 0% would be attained. Similarly, for attractiveness ratings, if a design was universally rated as attractive then its overall score would be 100%, and if universally rated as unattractive, then its overall score would be 0%. One subject failed to complete this part of the questionnaire, and so scores are calculated are from the remaining 39. The octolinear map scored 85.9% overall for the usability rating, and the concentric circles map 41.0%. The difference between designs was significant, $F(1,38) = 27.4$, $MSe = 1434$, $p < .01$. The usability ratings for each map were strongly and significantly correlated with the statement rating task aggregates. For the octolinear map, $r = .64$ and for the concentric circles map, $r = .64$, $p < .01$ for both, indicating consistent assessments of usability across different types of measure.

Interestingly, despite the answers given to individual items in the statement rating task, the gap between maps for the attractiveness rating was negligible. The octolinear map scored 65.4% overall, and the concentric circles map 64.1%. The difference between scores is non-significant, $F(1,38) = 0.018$, $MSe = 1743$, $p > .05$. The correlation between the attractiveness ratings and the usability ratings is $r = .49$, $p < .01$ for the octolinear map and $r = .35$, $p < .05$ for the concentric circles map. The correlation between attractiveness ratings and the statement rating task aggregates for the octolinear map is $r = .28$, $p > .05$ (non-significant) and for the concentric circles map, $r = .46$, $p < .01$. Overall, measuring opinions on usability, either via statement rating task aggregates, or via a single direct question, gives values that favour the octolinear map, and are consistent with each other. Attractiveness ratings favour neither map, but are related to opinions on usability, albeit not nearly as strongly, suggesting separate factors in play when people assess attractiveness versus usability.

3.3 The Relationship Between Subjective and Objective Measures

Thus far, the octolinear design has proven easier to use from the point of view of journey planning performance, and has received positive ratings. However, taking a closer look at these measures, it is clear that objective performance is not informing subjective ratings. In other words, the reasons why subjects are giving adverse ratings to the concentric circles map are not directly related to their actual performance at using the design. For example, there is no correlation between journey planning times and statement rating task aggregates for the octolinear map, $r = -.04$, $p > .05$, or for the concentric circles map, $r = -.14$, $p > .05$. This implies that subjects are rating usability according to factors that are not directly related to journey planning time, but by coincidence, for this particular pair of maps, the factors that lead to an adverse rating are simultaneously present with factors that lead to adverse usability. This need not be the case for other types of map.

Another way of looking at the relationship between subjective and objective measures is to look at people's map choice. If individuals are sensitive to one map being harder to use than another, then we would expect people to choose the map that was the easier for them to use. In other words, people who choose the octolinear map should have better journey planning performance for the octolinear map than the concentric circles map. People who choose the concentric circles map should have a better journey planning performance for the concentric circles map than the octolinear map. Although only six people chose the concentric circles map, this is sufficient to test this hypothesis. Table 1 shows that there is little evidence in its support. For people who chose the octolinear map, this design has a 6 second advantage on average. For people who chose the concentric circles map, there is still a planning time advantage for the octolinear map, albeit just under 5 seconds. The statistical test of this hypothesis is to conduct a two-factor mixed design Analysis of Variance with Map Choice as the between-subjects factor, and Map Type as the within-subjects factor, and see whether the interaction between the two factors is statistically significant. This was not the case, $F(1,37) = 0.294$, $MSe = 15.0$, $p > .05$. There is a possible interesting effect suggested by the means on Table 1: On average, people who choose the concentric circles map might be slightly faster at planning journeys in general than people who choose the octolinear map. However, this effect is not significant, $F(1,37) = 1.18$, $MSe = 128$, $p > .05$, meaning that replication with a larger sample size would be necessary before it could be concluded that this is a genuine effect.

People's map choices are not entirely arbitrary. Looking at Table 1, from statement rating task aggregates, it is clear that people who chose the concentric circles map rated this more highly than the octolinear map (44.0 versus 37.7). The reverse was the case for people who chose the octolinear map, who rated this more highly than the concentric circles map (45.4 versus 31.2). This time, using a two-factor mixed design Analysis of Variance with Map Choice as the between-subjects factor, and Map Type as the within-subjects factor, the interaction between the two was statistically significant, $F(1,37) = 17.9$, $MSe = 59.8$, $p < .01$. Hence, people's map choices are related to questionnaire ratings. Overall, people form strong opinions about the designs while using them, although not related to their performance, and these opinions have real consequences for whether a design is accepted or rejected.

3.4 Other Effects

Four items on the statement rating task were not directly related to aspects of map usability. Three of these concerned self-reported strategic approaches to journey planning (a preference for direct routes, avoiding complex areas on the map, and making as few interchanges as possible). The two maps did not differ significantly for any of these. In conjunction with the almost identical estimated journey duration measures for the two versions (see Table 1) these suggest that the different designs were not leading to different journey planning strategies. There was a significant difference for the question concerning geographical accuracy (mean score 3.7, *SD* 1.1 for the octolinear map; 3.0, *SD* 1.3 for the concentric circles map), $F(1,39) = 7.18$, $MSe = 1.36$, $p < .05$ indicating that subjects believed that the octolinear design was the more accurate of the two.

Next it was investigated whether age and self-rated expertise -- either at using rail transport in general, or direct knowledge of Berlin -- might be related either to journey planning performance, or to a subjective assessment of the maps via the statement rating task. The expertise measures are ordinal rather than parametric, but positive correlations would still be predicted. For example, people who reported more frequent visits to Berlin might be expected to plan journeys faster.

There were no significant correlations between age and journey planning time for either the concentric circles map or the octolinear version. The greatest correlation was for journey planning time for the concentric circles map, $r = .29$, $p > .05$. The direction indicates that older people had a slight tendency to be slower at planning journeys with this map (the octolinear correlation was smaller, but in the same direction, $r = .17$, $p > .05$). The statement rating task aggregate correlations were even smaller, indicating no relationship between age and map evaluation. Roberts *et al*. (2013) likewise failed to find any clear relationships between age and performance. The maximum age in the current study was 50 years old, and it is likely that more extreme values need to be sampled to identify clear evidence of age effects. Expertise correlations were similarly absent. The greatest value, $r = -.15$, $p > .05$, for the correlation between reported frequency of rail use and journey planning time for the concentric circles map, indicating slightly better journey planning performance for this map for more experienced uses, but was non-significant.

Turning to sex differences, it is a widespread belief that females are poorer than males at spatial tasks generally, such as navigation and interpreting maps, although Roberts *et al*. (2013) failed to find any trace of such effects. In the current study, there was no trace of sex differences in planning performance or usability ratings. However, for the attractiveness rating question, massive differences were found. Females rated the maps in general as being less attractive (mean rating 58.8%) than males did (68.4%). Using a two-factor mixed design Analysis of Variances with Sex as the between-subjects factor, and Map Type as the within-subjects factor, this main effect was significant, $F(1,37) = 5.32$, $MSe = 810$, $p < .05$, but there was also a substantial significant interaction, $F(1,37) = 7.82$, $MSe = 1478$, $p < .01$. The interaction indicates that males favoured the concentric circles map (a mean rating of 84.2%, compared with 60.1% for the octolinear version) but that females favoured the octolinear map (a mean rating of 70.0%, compared with 45.0% for the concentric circles version). Hence, although males and females did not differ in map usability ratings, there was a considerable difference in their assessments of relative map attractiveness.

4. Discussion

There are a number of clear findings from this experiment. Most importantly, the conventional octolinear map outperformed the concentric circles design both in terms of an objective measure (journey planning time) and a variety of subjective measures. If we accept that the concentric circles map has a high level of coherence compared with the octolinear design, we can conclude that this is insufficient to compensate for the poor simplicity of its line trajectories. From the point of view of making prescriptions for designers of schematic maps, it would be reasonable to suggest that the first priority should be line trajectory simplification, attempting to improve coherence after this has been achieved. The positive ratings gathered informally for concentric circles designs appear to result from cursory impressions from brief overviews rather than actual attempts to use such maps, perhaps from male commentators, given the massive sex difference in attractiveness ratings found in the current study.

The results in this experiment might imply a close agreement between objective measures and subjective ratings of usability, but a closer look at the data suggests that, despite all subjects planning using both maps, the two types of measure are still uncorrelated. Those aspects of the maps that evoke *opinions* on them seem to be independent of those aspects that determine *usability*. By coincidence, these two aspects were unusually co-related for the particular maps in the current research. This corroborates and strengthens the findings by Roberts *et al*. (2013), whose subjects only planned journeys using single designs. Specifically, subjects seem unwilling or unable to take account of journey planning performance when rating maps, usually leading to a dissociation between objective and subjective measures.

A perennial problem for human factors research and usability studies is that generalizable findings are sought, but these can only be investigated by testing specific instances (Roberts, 2014a). The current research prompts us to reject the specific concentric circles map tested, and to favour the particular octolinear design instead, but this does not allow us to reject the use of concentric circles maps in all cases, or even for the city of Berlin. For example, it might be the case that although performing poorly initially, the coherence of the concentric circles map promotes better network learning than the octolinear one, so that with long term use its relative performance improves. It is also possible that the design here could be improved in terms of simplicity of line trajectories, again bringing its performance into line with the octolinear version. More generally, the concentric circles concept could be wrong for Berlin, but more effective for other cities with different network structures. For example, the Moscow Metro has a circle line, but the remaining routes are predominantly radial, crossing the centre of the city cleanly from suburb to suburb (Figure 6). Alternatively, although the New York Subway lacks any clear orbital line, a point of radiation just south of the tip of Manhattan permits a strong set of lines running the length of Manhattan, providing a firm foundation to the map (Figure 7). The key criterion, in all cases, is the extent to which the structure of the network is compatible with the design rules selected. In general, concentric circles designs present powerful images, which could encourage map engagement and more frequent use of public transport facilities for people whose aesthetic tastes are particularly resonant with this way of presenting information. Individual differences in schematic map preference are substantial, and perhaps could and should be accommodated by transport undertakings, by giving users a choice of design.

Figure 6. Concentric Circles map of the Moscow Metro.

Figure 7. Concentric Circles map of the New York Subway.

The failure of the concentric circles map in the current study should also not be over-generalized by concluding that all 'non-standard' approaches to schematic mapping are flawed, and that octolinearity really is a gold standard after all. This was disproved by Roberts *et al*. (2013). In general, the design rules for a map should be chosen so that, in conjunction with the network structure, the objectives of an effective schematic map, *simplicity, coherence, balance, harmony,* and *topographicity* can be best achieved. This may require experimentation with different levels of linearity, or the creation of entire design sequences (Roberts, 2012). At the very least, such a process will heighten the awareness of the designer as to the different solutions possible for creating a schematic map for a particular network, and the unique troublespots and difficulties that the network may present. Even so, it is clear from the current results that usability studies are required before any controversial design is put into production.

Objective usability testing will continue to be the needed until either of the following are satisfied. *First*, until we possess theories of design, cognitive planning processes, and measurement of usability derived from an analysis of the maps themselves, such that design may be guided and successful predictions made concerning usability. This could take the form, for example, of various algorithms for measuring different aspects of a schematic map, and these being combined into a single effectiveness measure so that different designs can be rank ordered. If this could then be shown to be correlated with actual measures of usability, then there would be no need for extensive testing in the future. Measurement of design parameters is an integral component for researchers who are attempting to create computer algorithms that automate schematic map design (e.g., Nöllenburg & Wolff, 2011; Stott, Rodgers, Martinez-Ovando, & Walker, 2011) and so this grand objective is not completely far-fetched. *Second*, objective usability testing could be dispensed with if it is possible to identify methods of gathering subjective ratings which are highly correlated with objective performance measures. This is particularly important because of cases such as Roberts *et al.* (2013), where the most usable map was chosen by only 50% of subjects. Dissociations between objective measures and subjective ratings are commonplace in this domain. The problem is that usability studies can be time consuming to perform and require large sample sizes. User ratings are quicker to gather, and in many cases are the preferred means of design evaluation. They are real, in the sense that they do determine the acceptability of designs, but so far it has proved difficult to identify the underlying factors of design which drive reactions and responses, and to disentangle the different dimensions of perceived usability and attractiveness.

The questionnaire (via questions 16 and 17) gave subjects an opportunity to make brief written comments on the designs concerning the most liked and disliked features of each. The caveats identified earlier concerning difficulties with people's awareness of their performance and cognitive processes would be expected to apply equally to written responses, but these could conceivably point towards aspects of the designs that were particularly salient to subjects, and drove their overall assessments. Their responses were placed into broad categories, including *attractiveness*, *clarity/complexity* (e.g., a simple statement that a particular design was clear, or an individual aspect of it, such as straightness of line trajectories, or navigability of interchanges), *familiarity*, *geographical accuracy*, *orientation/ organization* (statements tending towards a more global evaluation of the coherence of the design), and *usability* (unqualified statements simply stating that a design was easy or difficult to use). Overall, the octolinear map was overwhelmingly liked for its clarity/(lack of) complexity (17 statements), and the familiarity of its design rules (11 statements) but the most frequently stated dislikes were also concerning (lack of) clarity/complexity (7 statements) plus (dis)orientation/(dis)organization (5 statements). Seventeen subjects declined to identify any disliked aspect. As might be expected for an unusual design, responses to the concentric circles map were less consistent. Nine statements indicated a liking of its clarity/(lack of) complexity, five its attractiveness, and five its orientation/organization. Dislikes were more consistent, with nine statements alluding to the perceived (lack of) clarity/complexity of the map, and fourteen to its (dis)orientation/(dis)organization. Frequent comments were that the design was disorientating to look at, and the predominance of circles gave the impression that journeys would be roundabout and indirect. This final aspect could be the key to the statement rating task aggregates being uncorrelated with journey planning times, and hence objective and subjective measures being dissociated. Some subjects may have perceived that the routes that they were planning appeared to be indirect on the concentric circles map and rated it poorly for usability, but direct on the octolinear map and rated it highly. The people with this perception need not be those with poor journey planning performance. Hence, for this design, objective and subjective measures are *both* adverse and yet *also* independent.

A better understanding of these issues might mean that in the future it could be possible, with properly designed instruments, to obtain subjective assessments that are more closely aligned with actual usability, although expectations and prejudices are strong (Roberts 2014b) and difficult to overcome, which means that unusual designs will always have a rating disadvantage. Individual differences are also a perennial problem; should they be ignored, or accommodated? Irrespective of whether better models of usability, or better rating methods can be developed, the crucial message is that there is a need for evidence-based design. Whenever it is intended to convey information, whether via typography, or for wayfinding, in schematic mapping, or in the design of pictograms, then if there is the potential for the configuration of information to determine performance, then objective usability testing remains essential, and cannot be replaced with subjective ratings. Certainly, no selection decision should be made solely on the basis of mass popularity, or a public vote, as recently took place in Boston, USA (Boston Globe, 2013).

Finally, it is important to comment on the methodology used in this study, as certain aspects of it could be criticized. The *first* potential complaint is that by testing UK university students, the subjects were somehow special, and by being unrepresentative, their data are unrepresentative. This might have been a more valid criticism thirty years ago when fewer school leavers went to university and universities were more selective. For this complaint to carry any weight, it would have to be argued that some (unspecified) quality of this sample rendered the concentric circles map uniquely difficult for them, and that this design would yield excellent performance if tested on the general population. There are no good reasons to expect difficulty reversals in university samples versus the general population, nor to expect different patterns of data from questionnaire ratings. The *second* potential criticism is that real journeys were not planned and undertaken, so that a planning task, such as the one implemented here, yields different behaviour compared with actual map use. Again, a convincing reason would be necessary to explain how this could result in difficulty reversals in experimental versus field situations, and it would be a brave transport undertaking that put the concentric circles map in production in the light of these data, and a wealthy one that decided to commission further field studies despite them. The *third* potential criticism is that just one objective measure of performance was taken: times taken to plan complex journeys (although estimated journey duration was taken account of, and there was the potential for subjects to make journey planning errors). This criticism is acknowledged. There are other aspects of map use that should be investigated, such as station finding and network learning. Other methodologies include analyses of the perceptual properties of schematic maps (e.g., Rosenholtz, 2011) and use of gaze tracking methodology (e.g., Burch *et al*. 2014). Clearly, diverse measures of performance in relation to map design are highly desirable, so as to obtain convergent measures of usability, and to obtain a fuller picture of how users engage with these ubiquitous and yet widely misunderstood methods of visual information communication.

Footnote

1. The key to designing a concentric circles map is to identify the central point of radiation. A poor choice will result in an unbalanced design, and difficulty in configuring radial elements. For Berlin, irrespective of the design rules, the main problems in configuring the design come from a dense sequence of stations on U2 from Stadtmitte to Alexanderplatz, and projected stations on U5 from Alexanderplatz to Unter den Linden. To the west of this, the need to place dense stations on U1/2/3 from Gleisdreieck to Wittenbergplatz restricts possible solutions. A central point of radiation corresponding to *Stadtmitte* permits sufficient space to configure both difficult regions; Other centres cause one or the other to become over-compressed. Once these two areas are configured, the construction of the remainder of the map can flow outward from them, attempting to keep line trajectories simple and the design balanced and compact. Accommodating station labels is the prime determinant of circle radius. A decision must be made as to whether each element is radial or orbital

The octolinear map was created as part of a sequential analysis performed by Roberts (2012) in which, by systematically exploring design rules, the theoretical minimum number of corners for each line for this network was established. U1 was chosen to be a straight line to ground the design with a horizon, and S1/2/25 to give a clear vertical axis. The limitations for any linear design arise when actual line trajectories do not match the angles available, so that zig-zags are required in order to correct them. On the octolinear map, this is evident on U9 between Zoologischer Garten and Westhafen.

The structure of the Berlin network means that a compact legible schematic map without topographical distortion is impossible, and even more distortion results from the design priorities applied to the two creations here, but an effort was made to ensure that relative positions of nearby stations were not distorted.

References

Bartram, D. J. (1980). Comprehending spatial information: The relative efficiency of different methods of presenting information about bus routes. *Applied Cognitive Psychology*, *65*, 103-110.

Bornstein, R. F. (1989). Exposure and affect: Overview and meta-analysis of research, 1968--1987. *Psychological Bulletin*, *106*, 265-289.

Boston Globe (2013). MBTA map-making contest garners 17,000 votes. http://www.bostonglobe.com/metro/2013/09/21/map-making-contest-garners-votes/egeCO7x7Q8rbfciGVHaSyI/story.html [accessed 20/11/2015]

Bronzaft, A. L., & Dobrow, S. B. (1984). Improving transit information systems. *Journal of Environmental Systems*, *13*, 365-376.

Bronzaft, A. L., Dobrow, S. B., & O'Hanlon, T. J. (1976). Spatial orientation in a subway system. *Environment & Behavior*, *8*, 575-594.

Burch, M., Kurzhals, K., Raschke, M., Blascheck, T., & Weiskopf, D. (2014). How do people read metro maps? An eye tracking study. Schematic Mapping Workshop 2014, University of Essex, April. https://sites.google.com/site/schematicmapping/Burch-EyeTracking.pdf [accessed 20/11/2015]

Carpenter, P. A. Just, M. A., and Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices test. *Psychological Review*, *97*, 404-431.

Chabris, C., & Simons, D. (2010). *The invisible gorilla*. New York: Crown Publishing.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, *5*, 121-152.

Creative Review (2013). A circular subway map for NYC. http://www.creativereview.co.uk/cr-blog/2013/august/a-circular-subway-map-for-nyc/ [accessed 20/11/2015]

Die Welt (2013). Ein Psychologe macht S- und U-Bahn Pläne rund. http://www.welt.de/gesundheit/psychologie/article120467042/Ein-Psychologe-macht-S-und-U-Bahn-Plaene-rund.html [accessed 20/11/2015]

Die Zeit (2013). Menschen mögen Muster. http://www.zeit.de/2013/37/u-bahn-plaene-design-maxwell-roberts [accessed 20/11/2015].

Dow, A. (2005). *Telling the passenger where to get off*. Harrow Weald, UK: Capital Transport Publishing.

Fink, M., Lechner, M., & Wolff, A. (2014). Concentric metro maps. Schematic Mapping Workshop 2014, University of Essex, April. https://sites.google.com/site/schematicmapping/Fink-Concentric.pdf [accessed 20/11/2015]

Garland, K. (1994). *Mr Beck's Underground map*. Harrow Weald, UK: Capital Transport Publishing.

Gothamist (2013). New NYC subway map elegantly inspired by concentric circles. http://gothamist.com/2013/07/25/fun_map_nyc_subway_map_as_a_circle.php [accessed 20/11/2015]

Guo, Z. (2011). Mind the map! The impact of transit maps on travel decisions in public transit. *Transportation Research Part A*, *45*, 625-639.

Guo, Z., & Wilson, N. H. M. (2011). Assessing the cost of transfer inconvenience in public transport systems: A case study of the London Underground. *Transportation Research Part A*, *45*, 91-104.

ITV News (2013). New London tube map goes round in circles. http://www.itv.com/news/2013-01-30/new-london-tube-map-goes-round-in-circles/ [accessed 20/11/2015]

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it. *Journal of Personality & Social Psychology*, *77*, 1121-1134.

Leder, H., Tinio, P. P. L., & Bar, M. (2011). Emotional valence modulates the preference for curved objects. *Perception*, *40*, 649-655.

Meo, M., Roberts, M. J. and Marucci, F. S. (2007). Element salience as a predictor of item difficulty for Raven's Progressive Matrices. *Intelligence*, *35*, 359-368.

Metro (2013). New version of London Underground map shows circles are the war forward. http://metro.co.uk/2013/01/30/new-version-of-london-underground-map-shows-circles-are-the-way-forward-3374600/ [accessed 20/11/2015].

Newton, E. J., & Roberts, M. J. (2009). Applying cognitive psychology principles to Underground map design. Poster presented at London South Bank University research day, July.

Nöllenburg, M. (2014). A survey on automated metro map layout methods. Schematic Mapping Workshop 2014, University of Essex, April. https://sites.google.com/site/schematicmapping/Nöllenburg_survey.pdf [accessed 20/11/2015]

Nöllenburg, M., & Wolff, A. (2011). Drawing and labeling high-quality metro maps by mixed-integer programming. *IEEE Transactions on Visualization and Computer Graphics*, *17*, 626-641.

Ovenden, M. (2003). *Metro maps of the world* (1st ed). Harrow Weald, UK: Capital Transport Publishing.

Ovenden, M. (2005). *Metro maps of the world* (2nd ed). Harrow Weald, UK: Capital Transport Publishing.

Ovenden, M. (2007). *Transit maps of the world*. New York: Penguin

Ovenden, M. (2008). *Paris Metro style in map and station design*. Harrow Weald, UK: Capital Transport Publishing.

Ovenden, M. (2015). *Transit maps of the world* (2nd ed). New York: Penguin

Palmer, S. E., Schloss, K. B., & Sammartino, J. (2013). Visual aesthetics and human preference. *Annual Review of Psychology*, *64*, 77-107.

Primi, R. (2002). Complexity of geometric inductive reasoning tasks, contribution to the understanding of fluid intelligence. *Intelligence*, *30*, 41–70.

Roberts, M. J. (2005). *Underground maps after Beck*. Harrow Weald, UK: Capital Transport Publishing.

Roberts, M. J. (2012). *Underground maps unravelled, explorations in information design*. Wivenhoe, UK: Published by the author.

Roberts, M.J. (2014a). Schematic maps in the laboratory. Schematic Mapping Workshop 2014, University of Essex, April. http://www.tubemapcentral.com/articles/roberts_empirical_2014.pdf [accessed 20/11/2015]

Roberts, M.J. (2014b). What's your theory of effective schematic map design? Schematic Mapping Workshop 2014, University of Essex, April. http://www.tubemapcentral.com/articles/roberts_theoretical_2014.pdf [accessed 20/11/2015]

Roberts, M.J., Newton, E.J., Lagattolla, F.D., Hughes, S., & Hasler, M.C. (2013). Objective versus subjective measures of Paris Metro map usability: Investigating traditional octolinear versus all-curves schematic maps. *International Journal of Human Computer Studies*, *71*, 363-386.

Roberts, M. J., Welfare, H., Livermore, D. P. and Theadom, A. M. (2000). Context, visual salience, and inductive reasoning, *Thinking and Reasoning*, *6*, 349-374.

Rosenholtz, R. (2011). What your visual system sees when you are not looking. *Proceedings of the SPIE: Human vision and electronic imaging*, *XVI*, San Francisco, California.

Silvia, P. J., & Barona, C. M. (2009). Do people prefer curved objects? Angularity, expertise, and aesthetic preference. *Empirical Studies of the Arts*, *27*, 25-42.

Stott, J. M., Rodgers, P. J., Martinez-Ovando, J. C., & Walker, S. G. (2011). Automatic metro map layout using multicriteria optimization. *IEEE Transactions on Visualization and Computer Graphics*, *17*, 101-114.

Vertesi, J. (2008). Mind the Gap: The London Underground map and users' representations of urban space. *Social Studies of Science*, *38*, 7-33.

Vodegel Matzen, L. B., Van der Molen, M. W., & Dudink, A. C. (1994). Error analysis of Raven test performance. *Personality and Individual Differences*, *16*, 433-445.

Wolff, A. (2007). Drawing subway maps: A survey. *Informatik*, *22*, 23-44.