This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier’s archiving and manuscript policies are encouraged to visit:

http://www.elsevier.com/copyright
Mind the map! The impact of transit maps on path choice in public transit

Zhan Guo

Robert F. Wagner Graduate School of Public Service, Rudin Center for Transportation Policy and Management, New York University, 295 Lafayette St., Room 3038, New York, NY 10012, USA

1. Introduction

Traveling in a transit system often involves a greater degree of uncertainty than traveling on a road network due to the complexity of transit systems and the stochastic nature of services (Hickman and Wilson, 1995). Transit users often need more information in planning travel than road travelers, such as operating hours, fare and fare media, waiting and travel times, access and egress, transfers, station locations, and seat availability (Abdel-Aty et al., 1996; Cluett et al., 2003). Such information is critical to passengers’ travel decisions (Khattak et al., 2003; Chorus et al., 2007), and the provision of information could be a powerful planning tool to guide individual decisions and to enhance the overall efficiency of the system (Polak and Jones, 1993; Ben-Elia and Shiftan, 2010).

While most studies in this field have focused on Advanced Traveler Information Systems (ATIS), this paper considers an alternative perspective and targets traditional information media. In particular, it focuses on the effect of schematic transit maps on travel decisions in public transit. The central argument is that a transit map has a tremendous impact on a passenger’s perceptions and his or her usage of the transit system. If implemented appropriately, a transit map can be a valuable tool to solve planning and operation problems in a cost-effective way.
This paper first develops a conceptual framework of the impact of the map on transit travel and then focuses on a specific travel decision: path choice in public transit. The case study of the London Underground indicates that passengers often (mis)trust a transit map more than their actual experience; they often take a path that looks shorter on the system map but is longer in reality compared with alternative paths. They also try to avoid transfer stations when the coded connection on a map looks less convenient than it actually is. The implications of the map on transit planning and operations are also discussed.

2. Literature

The importance of maps to spatial behavior has been well documented (Woods, 1992; Hutchins, 1995; MacEachren, 2004). However, in the field of transportation, maps have attracted little attention. Little is known about the travel information delivered by a map to a traveler and the effect of a map on travel decisions. There have been few efforts to incorporate the map as an analytical tool for transportation planning (Jankowski et al., 2001). The author identifies only a few related studies. Garland et al. (1979) tested the effect of transit map design on the quality of trip planning. They found that when the street detail was high, color coding of transit lines led to greater trip planning accuracy, less perceived difficulty, less frustration, and higher confidence. When color coding was absent, greater street detail led to lower trip planning accuracy, greater perceived difficulty, greater frustration, and lower confidence. Hall (1983) investigated students’ travel times by foot or bus from a university campus to a library 1.5 miles away. He found that those supplied with transit maps reached the destination significantly faster than those without maps did. However, students supplied with both schedules and maps were actually slower than those using maps alone. Dzieman (2008) found that a transit map is the primary source of information for passengers traveling to unfamiliar places.

Substantial progress has been made only very recently. Vertesi (2008) investigated the mental map of more than 20 London residents using the cognitive mapping approach. She found that the schematic underground map could significantly affect people’s spatial cognition and their wayfinding behavior in London. Hochmair (2009) compares the effect of four different transit map designs on route choices in Vienna. He found that a transit map with headway information allowed planning for faster routes compared with a transit map lacking this information. Both map annotation and network geometry affect the selection of the fastest route. Raveau et al. (2011) investigated the impact of the Metro map in Santiago, Chile on passengers’ route choice and found that route directness and angular cost depicted by the map matter. The model based on the map topology produced a slightly better goodness-of-fit than the model based on the actual topology, though the authors did not compare them statistically.

Three other types of literature are also relevant to the topic. The wayfinding literature generally investigates the effect of a map on people’s wayfinding behavior (Golledge, 1999; Devlin and Bernstein, 1997). Freksa (1999) compared maps and verbal route descriptions based on their effectiveness in assisting wayfinding. She also conceptually discussed an optimal trade-off between competing design criteria as well as seeking an adequate compromise between a “faithful” and a schematic map. Soh and Smith-Jackson (2004) looked at the influence of map designs, individual differences, and on-site signs on wayfinding performance in terms of the total time for completion, time and accuracy of decision-making, and deviation from the route. Other studies in this literature examined the difference between a map and the actual experience in spatial orientation (Tversky, 1981; Thorndyke and Hayes-Roth, 1982; Richardson, 1999; Uttal, 2000).

The cognitive mapping literature occasionally looks at travel decisions (Golledge and Garling, 2004; Weston and Handy, 2004). However, most focus on the effect of travel behavior on cognitive maps, not vice versa. For example, Mondschein et al. (2010) investigated the influence of mode choices on the formation of different cognitive maps of Los Angeles. Chorus and Timmermans (2010) compared the actual and perceived cognitive maps of passive (e.g., car passengers) and active travelers (e.g., car drivers or pedestrians) in Eindhoven, Belgium. Because of the difficulty quantifying cognitive maps, studies on the effect of cognitive maps on travel are generally qualitative (Hannes and Janssens, 2009) or use crudely simplified cognitive maps, such as perceived distances (Horning et al., 2008), a single spatial syntax value (Lee and Ryu, 2007), or frequency of travel (Arentze and Timmermans, 2005). This literature has not yet been applied to transportation planning (Golledge and Garling, 2001).

The third type of literature is cartography and geo-visualization that focuses on the quality of map design (Glasgow et al., 1995; Mijnsemaar, 1999). Some studies looked specifically at travel maps, such as the rise and fall of strip format travel maps by MacEachren and Johnson (1987), the differences in transit maps of France, Germany, and Spain (Morrison, 1994), or automated drawings of transit maps (Wolff, 2007). Berendt et al. (1998) proposed suitable criteria to evaluate the accuracy or inaccuracy of map-like diagrammatic representations and applied the criteria to develop a diagrammatic transit map of Hamburg, Germany. Only a few studies in this literature have explored the effectiveness of graphic displays for exploratory data analysis, problem solving, and learning (Shah and Miyake, 2005). For example, Fabrikant et al. (2010) studied users’ responses to different weather map designs in terms of eye movements and the time and accuracy of responses.

In summary, travel maps, in general, and transit maps, in particular, are still an unexplored multidisciplinary research topic. In particular, we know little about three critical questions on this topic:

1. What kind of information is delivered from a travel map to a traveler?
2. How could that information affect a traveler’s decisions?
3. Could we use a travel map as a planning tool to improve an individual’s decision-making and the performance of a transportation system? And how?

This paper covers the first two questions and leaves the third question for future research. Section 3 discusses these two issues conceptually and proposes a framework for subsequent empirical analysis. Because bus and urban rail systems are distinct in terms of map design, embedded travel information, and the possible effect on travel behavior, this paper differentiates the two and focuses on the urban rail systems.

3. Transit map, travel information, and travel decisions

3.1. Transit map and travel information

A transit map is a schematic diagram that depicts the locations, directions, and connections of stations and lines in a public transit system. It normally does not include service information, such as travel time or crowding. A transit map can deliver four types of travel information: distortion, restoration, codification, and cognition.

3.1.1. Distortion

A schematic transit map is often geographically inaccurate in terms of the depicted distance and direction. It may display the distance between two stations as either shorter or longer than it actually is. A line depicted as straight on the map may actually have many changes in direction. Most schematic maps use a fixed distance between stations, compress the outer area of the system, and expand the center. Such distortion and simplification can ease the burdens of information processing and facilitate passengers’ decision-making (Larkin and Simon, 1987; Hochmair, 2009). Detailed geographical information might be neither necessary, especially to passengers traveling underground, nor feasible, especially to large systems with many lines and stations packed in a small area in the urban center.

Distortion often prevails in urban rail transit maps. The author measured the map and actual distances for each link in the London Underground and calculated their correlation (Fig. 1). The value is only 0.22. In other words, the London Underground map only represents about 4% of the variation of the actual spatial distance between stations. Famous examples include those between Bank and Moorgate stations and between Piccadilly Circus and Leicester Square stations. Both are much closer than they appear on the map: by walking it takes only about 6 min in the former case, and 4 min in the latter case.

---

2 One exception is the subway map in San Francisco in the late 1960s. For each station, it showed peak travel times in minutes from Embarcadero or 12th and Broadway and marked the number on each station (Ovenden, 2007).
Other examples include the Central and District lines between Paddington and Notting Hill Gate, which run east–west on the system map but physically run north–south. The Notting Hill Gate station is actually west, not south (as depicted by the map), of Paddington (Fig. 2).

Fig. 2. London Underground in central London.
3.1.2. Restoration

Restoration refers to retaining geographic features in this generally distorted map environment. A few system maps remain largely geographical, such as those in New York City, San Francisco, and Chicago. Some are hybrid, like those in Madrid, Amsterdam, and Mexico City. Such variation depends on several factors: the structure of the system, e.g., an overground structure tends to be more geographical on a map than an underground one (Chicago); the spatial layout of the city, e.g., those with clearer geographical features tend to have more geographic maps (New York City and San Francisco); or the unique history of the transit map development, e.g., in Madrid and Paris (Ovenden, 2007). Restoration also occurs because distortion has its limits; it is possible for passengers to find too much distortion unacceptable. The 1972 subway map in New York City was finally abandoned 9 years after its introduction because of this particular reason. For example, Central Park was drawn as a wide, west–east rectangle even though it is actually narrow and runs north–south. Also, the Number 1 line 50th Street station on Broadway is east of 8th Avenue, but on the map it was shown west. Because of the clear grid network, New Yorkers often know the geographic location of stations and lines, so many found this distortion confusing.

Restoration is normally done by including geographic features, such as major roads, green spaces, or mountains on the map. A prime example is the current New York City subway map, which overlaps with the street network with many stations named after street numbers. Some transit agencies provide geographic maps as an alternative to the schematic map to passengers. For example, in Paris the transit maps on station walls remain fully geographic. In Sao Paulo, Brazil, a geographic map with all of the subway lines in the region is printed on the reverse side of the diagrammatic map (Ovenden, 2007).

3.1.3. Codification

Codification refers to how lines, stations, and connections are coded as symbols on a schematic map. The effect of map coding on map use has been well documented in the cartography literature (Simon and Larkin, 1987; Montello, 2002). Colors, visibility, and labels make some stations and lines more prominent in people’s cognitive maps than others (Garland et al., 1979; Dziekan, 2008). Another good example is Mexico City, where each station on the subway map has its own emblem to help the illiterate (Ovenden, 2007).

Codification, in general, is essential to denoting connections on a transit map. There are diverse ways to represent transfers, such as overlapped stations (most systems), semi-overlapped stations (Kiev, Ukraine), separate stations connected by a link (Moscow, Berlin, Tokyo), or no connection at all (Bilbao). Within the case of overlapped stations, the symbol could vary by size (Munich), color, or pattern, e.g., segmented dots (Moscow, Mexico City, Budapest), national flag symbol (Seoul), or a scheme too complicated to interpret (Miami).

Different codifications of connections could likely affect passengers’ transfer decisions in public transit. For example, in the London Underground, the Piccadilly and Victoria lines intersect at two transfer stations: Green Park and King’s Cross. Such a connection is represented as a simple, overlapped dot at Green Park compared with a long link between two separate platforms at King’s Cross. However, in reality transferring at King’s Cross is more convenient than at Green Park. This may mislead transferring passengers to avoid King’s Cross and to “over-transfer” at Green Park. The same “error” occurs at two other transfer stations, Charing Cross and Embankment, on the Bakerloo and Northern lines (see Fig. 2).

3.1.4. Cognition

Cognition refers to the cognitive effect of a schematic transit map on the perception of the transit system and the whole of urban space. Human psychology has shown that the diagrammatic type of visualization affects internal cognitive representations (Uttal, 2000; Hegarty, 2004). A transit map certainly affects the perception of the system: not only the simplicity or complexity of the system structure but also the overall image of the transit agency matter. Many transit agencies view transit maps as a valuable tool to establish a corporate identity.

However, the impact of the transit map likely goes well beyond the transit system (Vertesi, 2008). Transportation networks often act as a backbone in people’s cognitive map (Lynch, 1960). A transit network, the subway in particular, could potentially reshape the mental map of an urban space in at least two ways. First, underground travel substitutes surface travel and “deprives” the passenger of a chance to form a spatial cognition of the region through their own experience. Second, a transit map often offers alternative “assistance” with a clear, simplified, stable, and widely-published depiction of the urban structure. Deprivation and assistance, acting as push-and-pull forces, often occur at the same time and place, reinforcing the impact of the transit map on spatial cognition.

Specifically, a transit map may affect three elements of a cognitive map: boundary, landmark, and perceived distance. Subway lines often act as boundaries of different areas in a metropolitan region. A circumferential line might become the unofficial definition of a downtown or urban center (e.g., London, Chicago, Moscow, Berlin, Beijing). When different fare zones are adopted, they often become a proxy for different real estate markets. Major subway stations often become new or reinforce existing landmarks in a cognitive map. Hannes et al. (2006) show that towns or neighborhoods named by subway lines are more prominent on people’s mental map. In some cases, the number of stations, rather than the number of miles or kilometers, is used by the public to measure distance. Some critics argue that schematic transit maps actually

---

3 Personal conversation with John Tauranac, chairman of the 1975 MTA map committee that ended the 1972 map design and the design chief of the 1979 subway map, on June 20th, 2010.
represent “an ideal image of modern time and space: orderly, lucid, regular, efficient, and entirely functional” for the era of capitalism (Hadlaw, 2003).

3.2. Transit map and travel decisions

The information delivered by a transit map through distortion, restoration, codification, and cognition might affect travel decisions in two ways. First, it influences the number of available opportunities and travel options perceived by travelers in terms of destination, mode, or path choices. Second, it affects the perceived attributes of these opportunities and travel options, such as the desirability of a place and travel time and cost (Chorus et al., 2007). Affected decisions include location, mode, and path choices.

3.2.1. Location choice

If a transit map can affect people’s spatial cognition, it may change their location preferences. For example, it is believed that Harry Beck’s famous design for the London Underground was initially accepted partly for the reason of promoting suburban developments by making them look closer on the map to the urban center (Garland, 1994; Wolmar, 2004). Otherwise, the true distance from the center to these communities, such as those in the northwestern part of London, would be formidable and discourage people from moving there. Also, locations on the map may appear more prominent in one’s cognitive map and achieve a higher desirability. Vertesi (2008) illustrated this “Here be dragons” attitude in London where “spaces not colonized by station dots or Tube lines present uncharted and perhaps dangerous territory.”

Although there have been no empirical analyses to quantify this effect on location choice, studies have shown that spatial cognition, which can be influenced by a transit map, affects various location choices, such as residential preferences (Johnston, 1972; Pacione, 1982) and business locations (Pacione, 1978). The perceived distance, which is often distorted by a transit map, could affect store choices for shopping (Garling, 1989) and recreational and leisure destinations (Golledge and Timmermanns, 1990).

3.2.2. Mode choice

A transit map could potentially advance one mode while penalizing another because of the codification and cognition effects. For example, the transit map for an urban rail system is often straightforward, easy to understand, and beautifully drawn while the bus map is often complex and confusing. It is difficult to determine the alignment of a bus route and the locations of stops and terminals from a bus map. When each bus route is printed on a separate map or brochure, it is hard to go through all the printouts to identify the best bus route. In this context, transit maps may actually have “switched” customers from bus to rail, promoting one while haranguing the other.

Another example is that the enlarged center of a schematic map could make the stations in urban centers appear further apart than they actually are, thus penalizing walking. Passengers could have left the system earlier and walked to a destination if they had known how close the distance is in reality. In order to correct this perception bias, one walking advocacy group in London redrew the underground map and marked the walking time between stations in Central London.4 Many Londoners were surprised by the short walking distances (Fig. 3). Walking could be potentially faster than the underground between some stations, especially in peak hours. Noteworthy examples include Queensway and Bayswater (northwest corner of Fig. 2), and Mansion House and Bank stations (southeast corner of Fig. 2), which are much closer to each other than the map depicts.

A transit map might also alter the “transaction cost” of a mode switch. For example, when the street network is excluded from a transit map, and transit stations are not shown on a street map, drivers and transit passengers may develop different knowledge sets or cognitive maps of the two transportation systems. This is likely to make mode switching between transit and automobiles more difficult. A similar situation might occur with the bus and rail system. Although there are no empirical analyses on this topic, the studies by Fujii et al. (2001) on mode switch during a freeway closure and by Chorus et al. (2007) on the mode-switch effect of transit information on car drivers illustrate this possibility.

3.2.3. Path choice

A transit map may affect a passenger’s path choice through distortion, codification, and cognition. The path choice issue arises when a passenger has to select the “best” path over multiple options from an origin to a destination. Path attributes that are crucial to path choice decisions include path distance, direction, landmarks, connection, and service quality.

Path distance (or travel time) has been proven to be the single most important factor in path choice decisions (Dijkstra, 1959; Gallo and Pallottino, 1988; Prato, 2009). Numerous studies have found that path direction is also important. People generally prefer routes that are initially long and straight even if these routes are not the shortest in terms of Euclidean distance (Bailenson et al., 2000). They may follow a “least-angle” strategy to minimize the total length of the (perceived) initial street segment and the fictitious segment running from the endpoint of the initial segment to the distal target (Hochmair and Frank, 2002). People often prefer fewer directional turns along a path (Golledge, 1995; Heye and Timpf, 2003), tend to conserve linearity along their paths (Dalton, 2003), and follow landmarks (e.g., major stations on a map) in wayfinding.

---

4 http://shortwalk.blog.co.uk/2007/02/09/think_outside_the_tube_download_central_~17,07,446/
even they may not be the most direct path (Vertesi, 2008). A path with better spatial cognition, as represented by a higher score on spatial syntax, is more likely to be chosen (Lee and Ryu, 2007). Transfer is always a major concern for public transit passengers because of the inconvenience and uncertainty involved (Guo and Wilson, 2004).

A schematic transit map is able to change all these path attributes through distortion, codification, and cognition, and unavoidably alter the relative attractiveness of available paths. Using the London Underground as an example, there are two alternative paths traveling from Paddington to Bond Street station, path 1 transferring at Becker Street and path 2 transferring at Notting Hill Gates (Fig. 4). Path 2 is about 15% slower by in-vehicle time than path 1, and the Notting Hill Gate station is to the opposite direction of the destination Street (Fig. 4b). We would expect that few passengers would choose path 2. However, more than 30% of passengers chose path 2, probably because, on the schematic tube map, path 2 is about 10% shorter than path 1, and Notting Hill Gate station is shown to the south not west of Paddington (Fig. 4a).

In summary, a transit map can deliver multiple layers of information to travelers, and this information can have a strong and lasting effect on various travel decisions. To test the assumption, a large-scale empirical analysis was conducted for one
of the largest urban rail systems in the world, the London Underground. It focused on passengers’ path choices because the map was more likely to impact path choice than mode or location choices. Path choice also tends to involve less external “noises” and therefore is more likely to capture the pure effect of a transit map.

4. London Underground and the tube map

London has the world’s oldest underground railway system with services operating since 1863. The system had only one line at the beginning, so the map was simple and fully geographic. When the system grew bigger and became more complex, geographical details began to impair the legibility of the map, and many stations in outer London were cut off from the map (Garland, 1994). Many experiments were performed in the early 1900s, which mainly focused on Central London and gradually teased out the geographical features, leading to the adoption of a simple diagram (Ovenden, 2007). The first diagram map appeared in 1916 but only for the Piccadilly Line. In 1926, the geographic background was first omitted from the map but the topographic reference to the physical layout of tracks was preserved (Garland, 1994). The first system-wide diagram map appeared in 1933, based on Harry Beck’s sketch in 1931 which set the standard and became the permanent system map, often called the tube map.

Inspired by the sewage system map, Beck made the map conform to a set of rules: multiple colors, fixed station spacing, and single angles. Curving lines were smoothed and the center was enlarged. There are only vertical, horizontal, or 45° lines on the map, and all geographic features except Thames River were removed. Beck’s design was flexible enough to incorporate major system expansions. For example, despite the addition of almost 110 stations and six lines, the 2007 tube map is still remarkably similar to the original 1933 map.

Beck’s map is hailed as a design classic. Each year, the London Underground prints fifteen million copies of the map and distributes them free to passengers,5 and the diagram is reproduced 60 million times by other companies and institutions.6 The design principles have been followed by many other transit systems worldwide, such as New York City, Berlin, Munich, Paris, Hong Kong, Moscow, and Madrid. To many Londoners, the map is the business card of the city. It appeared on the city logo when London vied to host the 2012 Olympic Games. In 2006, it was voted by the public as one of Britain’s top three favorite designs over the past century.7 Beck’s map is also visible in Londoners’ daily life. It is a theme frequently used in interior decoration, furniture design, clothing design, tattoo, gifts and souvenirs, toys and games.8 To Londoners, this map is not just a simple diagram, but their cognitive map of London (Ovenden, 2007; Vertesi, 2008) – it is the “real” London. For example, in an interesting experiment conducted by Vertesi (2008), Londoners often draw a map of city based on the tube map: London is either represented as an x/y axis of the Northern and Central lines or a collection of distinct localities (underground stations). One of her interviewees made the point explicitly:

“I probably know London better by Tube than I do above ground, because when I’m walking without a map and then I hit a Tube stop, then I know where I am. So I sort of live in an underground world.”

This strong influence of the tube map is also partly due to the unique layout of London. The city does not possess a spatial structure that offers residents a clear visual identity, such as the grid system in New York City or the radial system in Paris. The street network is chaotic and wayfinding can be extremely challenging even with the assistance of a street map (Vertesi, 2008). For example, the posterior hippocampi of London’s taxi drivers were found to be significantly larger than others due to their frequent wayfinding work in this challenging environment (Maguire et al., 2000).9 To many Londoners, the tube map offers a short-cut to help them form a cognitive map of the city. The consistent adoption of the same map design over a long period of time also helps. Not many cities are able to do that. For example, New York City has had five versions of subway maps (1939, 1959, 1967, 1972, and 1979) during the same period of time, not to mention the continuous revisions in between.10

5. Modeling the map effect on path choices

To test the transit map effect on passengers’ path choices, a reference is defined as the reality. If a transit map presents a distorted reality, then the key research question is what passengers trust more: the schematic map or their own experience.

In this paper, the map effect is operationalized into two ways: map distance and transfer connection. The former is the distance of a path measured from a transit map. A passenger is often able to infer this distance by reading the transit map and taking into account the angles and turns of line segments, or connections. Therefore, the map distance is not just a measure of distance; it also accounts for the general layout of the transit system on the transit map. The latter refers to how a transfer station is coded on a transit map. Two types of codifications are compared: the dot connection with two overlapped stations and the link connection between two separate transfer platforms. A transfer may look shorter or longer on a transit map than it actually is.

5 http://art.tfl.gov.uk/projects/detail/1118/
6 http://www.bbc.co.uk/dna/h2g2/A673517
7 http://www.tfl.gov.uk/static/corporate/media/newscentre/archive/3649.html
8 Such as the famous Tube Map Lego and this find-animal-figures-on-the-tube-map exercise: http://www.animalsontheunderground.com/the_animals.html
9 Hippocampus is an important component of human brain responsible for the long-term memory and spatial navigation (O’Keefe and Nadel, 1978).
10 http://www.nycsubway.org/maps/historical.html
Related to these two aspects, there are two research questions:

Q1: Does map distance, relative to actual distance, affect passengers’ path choices? If so, how does the map distance interact with the actual distance as represented by the actual travel time?

Q2: Does the codification of transfer stations affect the perceived transfer inconvenience at the station? If so, how and to what extent?

To answer the first question, map distance was included in path choice models with the significance level tested, and its influence compared with that of the actual travel time. The variation of the map-actual relationship across different passenger groups was examined to understand the interaction between the transit map and travel experience. For example, the relationship could potentially be different between passengers who were familiar with the system and passengers who were first-time users (Lotan, 1997).

To answer the second question, the effect of codification on the perceived transfer inconvenience at a station was measured. Because transfer is unique to individual transfer stations, a dummy variable for the codification type (e.g., dot or link) could be “contaminated” by station-specific attributes. Therefore, the codification effect on the transfer inconvenience was captured by comparing models with and without the map distance. A convenient transfer station with a short distance transfer walk may look “bad” on a map when the transfer connection is represented by a long link. Because the perceived map distance is partially affected by the length of the transfer link, including the map distance in models could partly capture the codification effect on the perceived transfer inconvenience.

5.1. Data and variables

The main dataset is the Rolling Origin and Destination Survey (RODS), conducted by Transport for London (TfL) or its predecessor organization London Transport from 1998 to 2005. RODS records travel paths including the access, transfer, and egress stations for more than 250,000 trips in the underground network.

There are 691 station-to-station links (two directions) for more than 300 stations in the underground network. The map distance of each link was measured manually in Adobe Photoshop, so the unit has no real meaning. Travel times were obtained from RailPlan, a network model developed by TfL for public transport users for the AM peak period in London (RailPlan Modeling User Guide, 2006). These include in-vehicle travel time, station entry/exit time (walking time between the station entrance and a platform), initial waiting time, and transfer walking and waiting times.

Another interesting variable is the number of stations along a path. It might act as either a map or a “reality” variable because a passenger may count the number of stations from a transit map in addition to mentally calculating the path distance. The variable is also a good proxy for the actual travel time. Including it in models makes a fair comparison between the map distance and in-vehicle time. Table 1 lists the descriptive statistics for all variables.

<table>
<thead>
<tr>
<th>Path attributes (min)</th>
<th>Average</th>
<th>Max.</th>
<th>Min.</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>In-vehicle RODS paths</td>
<td>18.1</td>
<td>55.5</td>
<td>1.5</td>
<td>8.4</td>
</tr>
<tr>
<td>Generated paths</td>
<td>21.3</td>
<td>88.6</td>
<td>1.5</td>
<td>10.3</td>
</tr>
<tr>
<td>Entry within station RODS paths</td>
<td>1.8</td>
<td>6.1</td>
<td>0.2</td>
<td>1.0</td>
</tr>
<tr>
<td>Generated paths</td>
<td>2.0</td>
<td>4.0</td>
<td>0.1</td>
<td>1.2</td>
</tr>
<tr>
<td>Exit within station RODS paths</td>
<td>2.6</td>
<td>6.6</td>
<td>1.1</td>
<td>0.9</td>
</tr>
<tr>
<td>Generated Paths</td>
<td>2.8</td>
<td>7.8</td>
<td>1.1</td>
<td>1.4</td>
</tr>
<tr>
<td>Initial waiting RODS paths</td>
<td>2.0</td>
<td>10.1</td>
<td>1.0</td>
<td>1.3</td>
</tr>
<tr>
<td>Generated paths</td>
<td>2.5</td>
<td>16.6</td>
<td>1.0</td>
<td>1.8</td>
</tr>
<tr>
<td>Map distance * RODS paths</td>
<td>19.7</td>
<td>54.8</td>
<td>1.9</td>
<td>8.4</td>
</tr>
<tr>
<td>Generated paths</td>
<td>24.9</td>
<td>105.6</td>
<td>1.9</td>
<td>11.8</td>
</tr>
<tr>
<td># of Stations RODS paths</td>
<td>9.9</td>
<td>28.0</td>
<td>2.0</td>
<td>4.1</td>
</tr>
<tr>
<td>Generated paths</td>
<td>12.0</td>
<td>43.0</td>
<td>2.0</td>
<td>5.4</td>
</tr>
<tr>
<td># of Transfer RODS paths</td>
<td>1.0</td>
<td>2.0</td>
<td>1.0</td>
<td>0.2</td>
</tr>
<tr>
<td>Generated paths</td>
<td>1.1</td>
<td>3.0</td>
<td>1.0</td>
<td>0.2</td>
</tr>
<tr>
<td>Transfer waiting time ** RODS paths</td>
<td>1.9</td>
<td>30.2</td>
<td>0.0</td>
<td>1.9</td>
</tr>
<tr>
<td>Generated paths</td>
<td>2.2</td>
<td>30.2</td>
<td>0.0</td>
<td>2.1</td>
</tr>
<tr>
<td>Transfer walking time ** RODS paths</td>
<td>1.8</td>
<td>10.1</td>
<td>0.0</td>
<td>1.2</td>
</tr>
<tr>
<td>Generated paths</td>
<td>2.4</td>
<td>13.7</td>
<td>0.0</td>
<td>2.2</td>
</tr>
</tbody>
</table>

* Photoshop map unit; N = 2771 for RODS paths, and 6481 for generated path.
** From N = 1044 transfer movement directions for both RODS and generated paths.
5.2. Path choice models

The most challenging task of modeling path choice is identifying the alternative paths considered by a passenger. Unfortunately, most travel surveys like RODS only reveal the final decision (a chosen path), not the available options (alternative paths). Analysts normally rely on commonly accepted rules to infer these alternative paths (Bovy and Stern, 1990). In this research, the labeling approach developed by Ben-Akiva et al. (1984) was used to generate the alternative path choices. A label defines a set of weights for travel attributes. A shortest path algorithm then finds a path that minimizes the generalized cost based on the weights. When one travel attribute is weighted very heavily, the algorithm will find a path with the least value of that attribute.

Not all RODS origin–destination (OD) pairs and paths can be used for path choice modeling. ODs with too few trips were excluded as well as paths with more than two transfers and a smaller share (<10%) of trips compared with other paths between the same OD because they did not reflect a typical decision. Finally, 9284 RODS OD pairs, corresponding to 13,925 RODS paths, were selected. Eighty-five labels were then applied to these OD pairs on two types of underground networks: the default RailPlan network and the map-based network. In the latter case, the network-link attribute was the map distance instead of the actual travel time. In the process, 77% of the 13,295 revealed RODS paths, representing more than 90% trips, were generated on the default network. This result was comparable to the findings in prior studies (Ramming, 2002; Fiorenzo-Catalano, 2005; Bekhor et al., 2006). The choice set generation process and the definition of each label were fully described in prior studies (Guo, 2008; Guo and Wilson, 2011).

OD pairs were further filtered based on the result; those with only a single path option and those with no RODS paths generated could not be used for path choice modeling and were excluded. The final path choice set included 2330 RODS ODs, corresponding to 2771 RODS paths and 18,894 RODS trips. For these OD pairs, 6481 path options were generated. In other words, on average, one RODS OD had 2.8 generated paths. Generated paths tended to have larger averages and standard deviations than those revealed RODS paths, which we expected (Table 1).

A multinomial logit model was adopted. The path overlap issue was partially controlled by including dummy variables for 17 major transfer stations in Central London because most overlapped segments started or ended at one of these stations. The choice set size varied from OD to OD, ranging from two to six. Two models were developed with (Base Model) and without (Map Model) the map attributes.

Since the choice set was generated based on a reasonable “guess,” the validity of the modeling result had to be tested before making interpretations. The dataset was split into two parts, each with half the observations from the original dataset. The split was based on odd and even numbers of the access station codes, so the process was random. Next, half of the dataset was used to estimate parameters using the Map Model specification. Then, the parameters were used to predict the path choice in the other half of the dataset. The average probability of a correct prediction was 80% for paths and 91% for trips, suggesting that the choice set generation procedure and the model specification worked well. The modeling results are summarized in Table 2.

6. Results and analysis

Most variables in both models were significant at the 5% level with expected signs. In the Base Model, the more a path had in terms of entry, exit, in-vehicle, initial waiting time, transfer walking or waiting time, the less likely that path was chosen by a passenger. In-vehicle time was perceived as more onerous than the initial waiting time (−0.55 vs. −0.36) probably due to the high frequency of service. Transfer walking was more onerous than entry and exit walking (−0.32 vs. −0.29) while transfer waiting was viewed as more convenient than initial waiting (−0.2 vs. −0.36). This probably reflects the different weights of long- and short-term decisions when waiting in general is not a concern for the high-frequency service.

The 17 transfer station dummy variables captured the perceived transfer inconvenience by each station in addition to the transfer walking and waiting times. Such inconvenience is determined by physical transfer environments, such as escalator availability and ease of wayfinding, and possibly the codification of transfer stations on the tube map. All other transfer stations acted as a base for comparison, and their inconvenience was captured by the variable: number of transfers. A positive coefficient of the dummy variable indicated a better perceived transfer environment (in addition to transfer walking and waiting) than the base. The Base Model revealed a large variation of perceived transfer inconvenience across the 17 stations. The worst transfer stations were generally the large, complex, National Rail terminal stations, including Waterloo, Paddington, and Euston. The best transfer stations were the simple, heavily-used stations, such as Earl’s Court, Bond St., Leicester Sq., Oxford Circus, and Victoria.

When the map attributes were included in the Map Model, the explanatory power of the model increased significantly from 0.579 to 0.604 (significant at the 0.0001 level). The changes in variable estimations answered the two research questions.

6.1. Map distance more influential than actual travel time

The Map Model confirmed that the tube map indeed has an impact on passengers’ path choices. The map distance variable was statistically significant at the 1% level with an expected sign. The longer a path appeared to be on the tube map, the less likely that path was chosen with all other factors being equal.

11 The algorithm used is Pathfinder in TransCAD 5.0 implemented as a GISDK script. It is available from the author per request.
Comparing the relative influences between the map distance and the actual travel time, the average elasticity was calculated for all 18,894 trips (Table 3). The map distance had an elasticity of \(-0.0416\), meaning that a 10-percent increase of path length on the tube map would reduce the chance of that path being chosen by 0.416%. The elasticity for the actual travel time was much smaller (absolute value) at \(-0.0416\). The ratio between the two was about 2.14, i.e., the tube map was about two times more influential than the actual travel time on a passenger’s path choice in the underground. To test the interaction between the tube map and actual experience, the model was estimated again for two distinct groups: frequent users who rode the underground at least 5 days per week and first-time users. The rationale is that when a passenger becomes more familiar with the system, she or he would be more likely to rely on her or his own travel experience instead of using external information, such as a transit map, to make decisions (Emmerink et al., 1996). The modeling results confirm this assumption. For frequent users, the map effect was reduced as indicated by the smaller elasticity of the map distance \(-0.0387\), a 7% decrease, and the reliance on actual travel experience increased, indicated by a higher elasticity of travel time \(-0.0229\), 15% increase. However, the map distance was still more influential than the actual travel time, as indicated by a decreased yet still greater than one map-travel time ratio \(1.69\). For first-time users, both map and experience effects remained almost unchanged.
The tube map was even more influential for older (60 year or older) and younger (16–25 years old) users with a higher map-travel time ratio (2.8). However, that change was primarily caused by the reduced influence of actual travel time while the map effect remained unchanged. The elderly might be more sensitive to the ease, rather than the speed, of travel. Both older and younger users might have fewer time constraints compared with middle-aged users, who are more likely to be commuters.

6.2. Codification of connection affecting perceived transfer convenience

The coefficients of several transfer station dummy variables changed significantly after the map attributes were included, suggesting the potential effect of codification on the perceived transfer inconvenience. Such inconvenience was measured as the ratio of coefficients between a transfer station dummy and entry/exit time, the time variable that remained the most stable in both the Base and Map Models. The ratio is the unique “penalty” associated with the transfer station in equivalence of entry/exit walking time.

The tube map makes two transfer stations, Baker St. and Embankment, look more inconvenient than in reality. The coefficient of the Baker St. station increased from $-0.512$ to $-0.298$ (42%) with the latter being insignificant at the 5% level. In other words, the link for the Baker St. station on the tube map increased the perceived transfer inconvenience from 1.1 ($-0.298/-0.281$) to 1.8 ($-0.512/-0.288$) min of entry/exit walking. A similar result occurred at the Bank/Monument station. The perceived transfer inconvenience increased from 1.5 to 2.2 (33%) min of entry/exit walking.

Including the map attributes also reduced the perceived transfer inconvenience for some transfer stations, especially Victoria and Oxford Circus. In other words, they appeared to look like more convenient locations to transfer than they really are. Transfer at each of these two stations was coded as a simple dot even though the actual transfer environment was relatively challenging. For example, both stations are the busiest in the underground and operate consistently at or over capacity. Oxford Circus has the largest transfer volume, more than 110,000 per weekday, while at Victoria closing of ticketing gates to reduce platform overcrowding has become a daily practice. The easy connection coded for the two stations on the tube map, in turn, might contribute to this problem.

The modeling result also enabled us to quantify the effect of the codification of a transfer connection. For example, the extra or precluded transfers can be calculated for these stations using the elasticity of entry/exit walking time ($-0.02$). Results show that the link codification for the Baker St. and Bank/Monument stations might prevent 216 and 147 transfers, respectively, on a typical weekday. The dot codification for the Victoria and Oxford Circus stations might attract 960 and 516 more transferring passengers on a typical weekday (see Table 4). Codification of connections on the tube map could likely affect the daily operation of the underground, especially at transfer stations at or over capacity.

7. Discussion and conclusion

This paper investigates the effect of schematic transit maps on travel decisions in public transit systems. The relationship might have significant implications for public transit operation and planning, but so far it has been largely overlooked by both academics and practitioners. The paper first defines four types of information delivered from a transit map: distortion, restoration, codification, and cognition, and then discusses their potential influence on travel location, mode, and path choices.

The case study on the London Underground confirms that a schematic transit map indeed affects passengers’ path choices. Moreover, the map effect is almost two times more influential than the actual travel time. In other words, underground passengers trust the tube map (two times) more than their own travel experience with the system. The map effect decreases when passengers become more familiar with the system but is still greater than the effect of the actual experience, even for passengers who use the underground 5 days or more per week.

The paper also shows that the codification of transfer connections is also important. Different codification can make a transfer look more or less convenient on a transit map than in reality, which will either decrease or increase the perceived transfer inconvenience for the corresponding stations. This paper observes both situations in the underground case study.

---

and quantifies this codification effect, in terms of the number of attracted or precluded transfers, for four major transfer stations: Baker St., Bank/Monument, Victoria, and Oxford Circus.

Of course, these results are only based on the London Underground, a unique case in many aspects. Few transit maps enjoy such public popularity as the tube map in London. Many transit maps include prominent geographical features, which dilute the map effect. Other systems have different past or present versions of their transit map, which precludes a lasting and stable map effect. Many metropolitan regions possess an easier-to-comprehend urban form than London, which could weaken the role of a transit map in the formation of a cognitive map. The subway map effect in New York City is probably different from that in London. Therefore, readers should be cautious about making generalizations.

If a transit map has an impact on travel decisions, what are the implications for transit operation and planning? First, if passengers trust a schematic map more than their own experience, all planning efforts aimed at changing travel behavior need to consider the map effect; otherwise, the effectiveness of those efforts might be weakened. For example, this map effect might partially explain why Advanced Traveler Information Systems (ATIS) often yields modest improvements in terms of travel time savings in public transit (Hickman and Wilson, 1995; Avineri and Prashker, 2006; Ben-Elia et al., 2008). Second, a transit map might cause certain operational problems. For example, it might unintentionally shift more passengers to a congested segment in the network and thus form a bottleneck. The overcrowding at the Victoria and Oxford stations and on the link between the King’s Cross and Old Street stations, which is much shorter on the tube map than in reality, are possible examples.

Accordingly, a transit map could potentially become a planning tool to solve operational problems and improve system efficiency. For example, link lengths could be revised, and transfer stations could be re-coded on a transit map in order to change passenger behavior and mitigate platform and train crowding..Annotations of waiting time or crowding for selected stations on the map might also be important (Hochmair, 2009). Clearly, this approach has its own limits: we could not re-draw a transit map however we pleased.

In terms of future trends, ATIS and alternative travel information channels, such as smart phones and the internet, might change the role of a transit map in mixed ways. On the one hand, they may weaken the transit map effect. For example, internet-based trip planners may recommend specific travel paths based on their actual attributes. On the other hand, they may strengthen the map influence as well. For example, a transit map might become more accessible to passengers through, for instance, smart phones or the internet. Travel information, such as crowding and delays, delivered in a map format could be more effective than other media (Hato et al., 1999; Talaat, 2011). Conventional media like the transit map will still likely be critical and indispensable for trip planning despite the prevalence of real time information (Cluett et al., 2003).

In summary, transit maps can have a profound impact on passengers’ travel decisions and system performance. Both individual passengers and transit agencies should “mind the map” in order to make their best planning decisions.

Acknowledgements

The author wishes to thank Professor Nigel H.M. Wilson from MIT for his insight that made this study possible. He is also grateful for the valuable comments from Anthony Shorris, Rae Zimmerman, John Attanucci, Andrew Mondschein, and Amy Faust. The path choice model described in this paper was developed by the author with funding support from Transport for London (TfL) through the MIT-TfL Research Initiative.

References


