The Differentiating Behaviour of Shoppers

Ref 036

Clustering of Individual Movement Traces in a Supermarket

Jorge Gil
Space Syntax, NA, London, United Kingdom
j.gil@spacesyntax.com

Eime Tobari
Space Syntax, NA, London, United Kingdom
e.tobari@spacesyntax.com

Maia Lemlij
Space Syntax, NA, London, United Kingdom
m.lemlij@spacesyntax.com

Anna Rose
Space Syntax, NA, London, United Kingdom
a.rose@spacesyntax.com

Alan Penn
UCL, Bartlett School of Graduate Studies, London, United Kingdom
a.penn@ucl.ac.uk

Keywords
space syntax; supermarket; shopper behaviour; movement patterns; k-means clustering

Abstract
In this paper we report on research into patterns of shopper movement and behaviour in a supermarket. It is an underlying assumption of much space syntax research that the population is relatively homogeneous in terms of the way it uses space, introduced by the theory of natural movement in urban environments where the density distribution of facilities varies and the spatial configuration is diverse. However, in the specific case of a supermarket where goods are very evenly distributed in a regular and even grid, we need to explore the variations in individual behaviour within the population to understand the movement patterns in the store. Using data of over 480 shoppers interviewed and individually tracked in their shopping trip in a supermarket, we develop methods of profiling different shoppers according to both demographic and behavioural factors. We show that there are distinct clusters of shopping strategy defined in terms of characteristic search trail through the store, and that these correlate with specific shopper profile. We conclude that in situations where the allocation of attractors to space is neither random nor correlated to spatial structure, different groups of people on different kinds of mission will adopt distinctive spatial behaviour. This suggests that the assumption of homogeneity may be ill founded, at least under these circumstances.

1. Introduction

It is often assumed in space syntax studies that a population's use of space can be described by a simple correlation between spatial configuration and movement flow rates (Hillier et al. 1993, Peponis et al. 1989, Read 1999). This relationship appears to hold well for urban environments in which origins and destinations of movement are evenly distributed, and in which spatial configuration is relatively intelligible. It also appears to hold in urban environments in which origins and destinations are allocated to spatial configuration in an emergent manner such that density of attractors becomes correlated to spatial structure, as can be argued may be the case in long evolved and unregulated urban structures (Penn et al. 2004). Further studies have explored the social and economic impact of land use locations and densities in relation to accessibility and movement in urban environments (Marcus 2005, Stahle et al 2005, 2007). The underlying assumption and conclusion is that the population as a whole is relatively homogeneous in terms of its spatial needs as well as in its competence – for example, in its access to knowledge of the layout and location of destinations. However, in the specific case of supermarkets, where goods (attractors) are very evenly distributed in a regular and even grid, this assumption is open to question: surely real populations are made of many different groups of people each interested in different things at different times? Surely human society is much more complex and multi layered than the assumption of homogeneity would allow? Peponis et al's studies of detailed tracking of exhibition goers supports this more complex view (Peponis et al. 2003, 2004).

2. The behaviour of shoppers in a supermarket

In this paper we describe a study in which the assumption of the homogeneity of the population in its space use patterns is tested in a supermarket environment. Using data in which over 480 supermarket shoppers were individually tracked throughout their full shopping trip within a large supermarket we develop a method for segmenting the population into a set of groups defined by their differential characteristic space use patterns. We have set out the following key questions:

- How and to what extent does the spatial configuration of store layout impact shopper behaviour (e.g. spatial movement patterns, shopping duration and interaction with products)?
- Do any of the shopper groups express distinctive use of space or distinctive shopping behaviours?
- Can we identify distinctive movement patterns? And if so, are those patterns associated with certain shopper groups?

To answer these questions we used a detailed data set produced by a leading UK shopper research consultancy and the store's plan identifying the location of the different products. The data set is composed of shopper profile data obtained from interviews and shopper behaviour data extracted from video recordings, thus covering stated preference (Miller et al 1998) and revealed preference types of information for more comprehensive and unbiased results. The shoppers were initially approached to agree to take part in a survey, record basic information on age, gender, group size, carrier type and clothing and be given a coloured tab to identify them upon exit. This initial information was used by camera operators to follow the shoppers within the store using the store's CCTV system and record their every movement. Shoppers' behaviour was only marginally affected during the trip since the initial approach was very brief and the observation method was unobtrusive. On exiting the store an extensive interview was conducted covering aspects of the specific trip, e.g. aim of trip, use of shopping list, satisfaction, money spent, as well as more general shopping habits, e.g. frequency of visits. The video recordings were post-processed to extract a data set of the store areas visited, with a time stamp on entry and exit of each area, and the types of product interaction happening in each area, also with a time stamp at beginning and end of the interaction.

To analyse this data set we performed spatial and statistical analyses using MapInfo Professional GIS and Confeego by Stutz et al., UCL Depthmap by Turner and JMP statistical exploration software from SAS.

3. Testing the homogeneity assumption

Firstly, we ran spatial configuration analysis of the store layout in terms of accessibility and visibility and correlated this with shopper movement flows, to test the extent to which people's movement patterns and interactions are affected by the spatial layout alone.

Various types of space syntax models were used as is typical of research on complex buildings (Turner 2003, Koch 2005, Markhede and Koch 2007, Sailer 2007), including axial and segment models and visibility graphs, where graph, isovist and agent analyses were performed. However, due to the simple spatial configuration of the store, the level of detail of the shopper and product data sets, the visibility graph proved more relevant, offering higher resolution and unique spatial values for each of the store's data units (product sectors). For spatial accessibility we ran VGA analysis, where the store layout was divided into units of 30 by 30 centimetres, and calculated visual integration and distance from the entrance. The spatial structure of the store in terms of spatial accessibility is quite simple. As demonstrated in Figure 1, the spatial accessibility values can be categorised into three ranges based on natural break: low, medium and high.

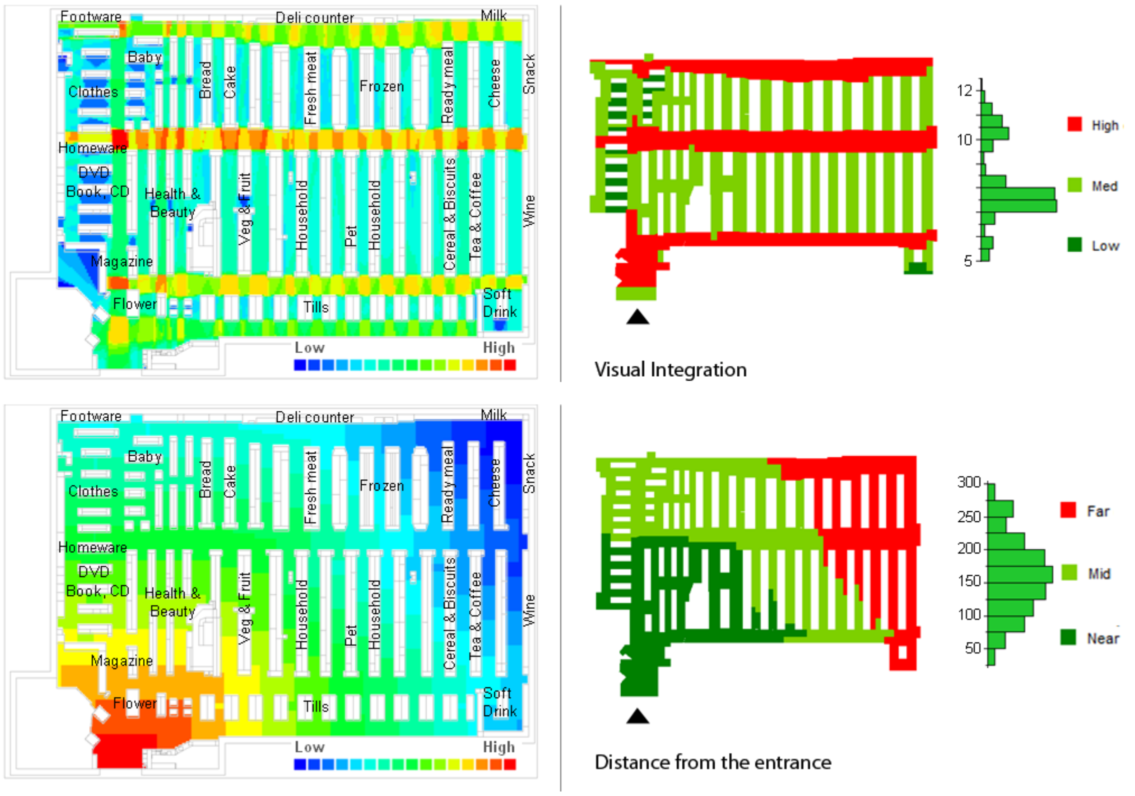

Figure 1
Spatial accessibility analysis of the store: visual integration (above) and distance from the entrance (below); simplified classification of the space on the right.

- the central corridor is the most accessible space within the whole store, followed by the two parallel corridors running to either side
- the entrance corridor and all aisles have similar values of spatial accessibility
- aisles in non-food sectors (CD, DVD, books, etc) are the most segregated spaces in the store despite their proximity to the entrance

After performing correlations between movement flows and spatial accessibility (Figure 2) we concluded that shopper behaviour is strongly affected by product locations, which is not a surprise since that is the main purpose of the shopping trip and the layout of the store is very simple.

A high level of movement was found around the entrance and in the main corridor, followed by the other two parallel corridors, which correlate with the spatial accessibility patterns of the store. In these areas there is very little interaction with products as they are mostly used for navigation and way finding. The lowest level of movement is found in the non-food and baby sectors, which also correlates with lower spatial accessibility.

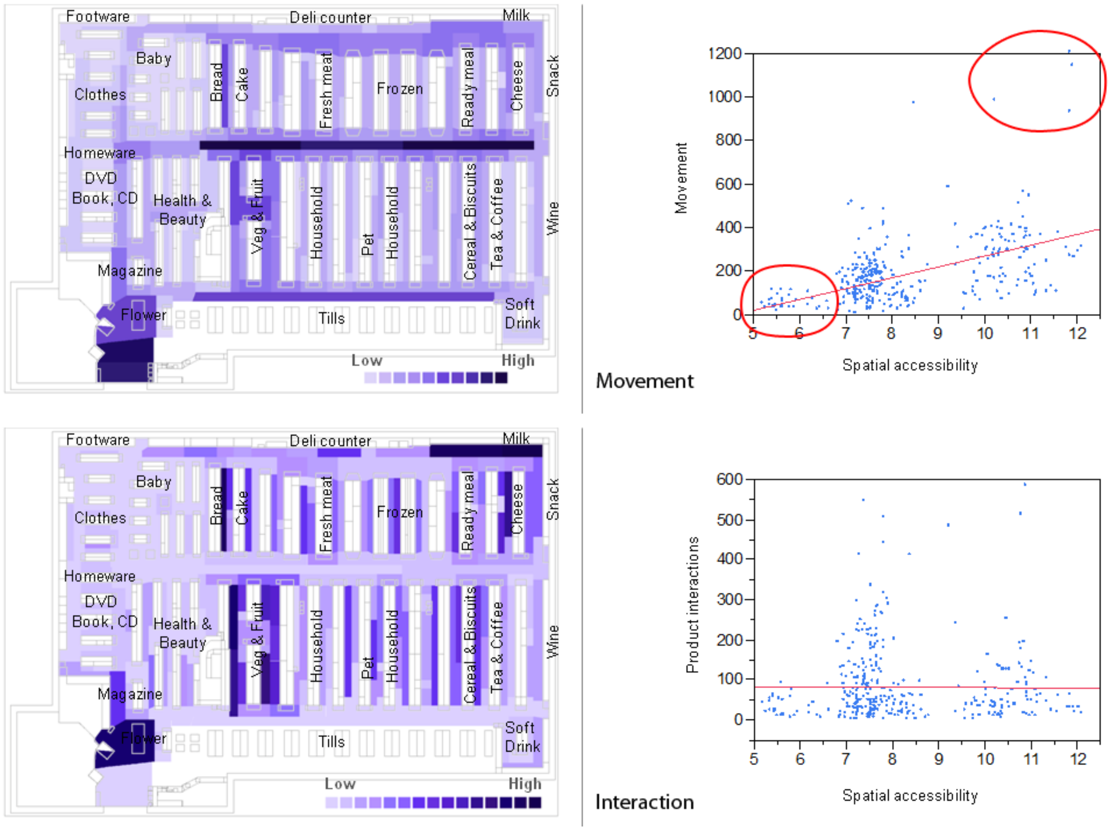

Figure 2

Correlation between spatial accessibility and shopper movement (top) is low (R^2 = 0.26) except in the main corridors and non-food and baby sectors (circled in red). There is virtually no correlation between spatial accessibility and level of product interactions (bottom), as most interactions happen in the aisles which have a very similar level of accessibility.

However, in the major part of the store, movement is driven by interactions with products, with the highest level of interaction coinciding with the popular product sectors, i.e. Milk, Bread and Fruit & Vegetables. There is virtually no correlation between spatial accessibility and the level of product interactions because in this part of the store there is a very similar level of spatial accessibility throughout. This results in a low overall correlation between the spatial accessibility of the store and shoppers' movement (R2 = 0.26).

The main conclusion is that given the current design of supermarkets there are two systems: one for movement and another for interaction with goods. The spatial configuration of store layout responds to some extent to the natural movement needs, but a different set of principles appears to affect shoppers' behaviour, with product location being the main factor. This conclusion seems to confirm findings in relation to office spaces presented by Sailer (2007). However a question remains: does this hold for every shopper, do they all show the same use of space?

4. Population filtering

We proceeded to examine the shopper population data in more detail using simple data filtering techniques. From both observations and questionnaires we selected a number of variables of shoppers' demographics, profile and behaviour. The selected variables were:

- Shopper demographics - Gender, age, household size, socio-economic group, employment status
- Shopper profile - Carrier type, group size, shopping mission, frequency of visit, shopping list, attitude towards promotions, method for finding their way around, age of the accompanied

- Shopping behaviour - Shopping duration, money spent, backtracking rate
- Atypical groups - People who purchase baby products, alcohol, milk, pet care

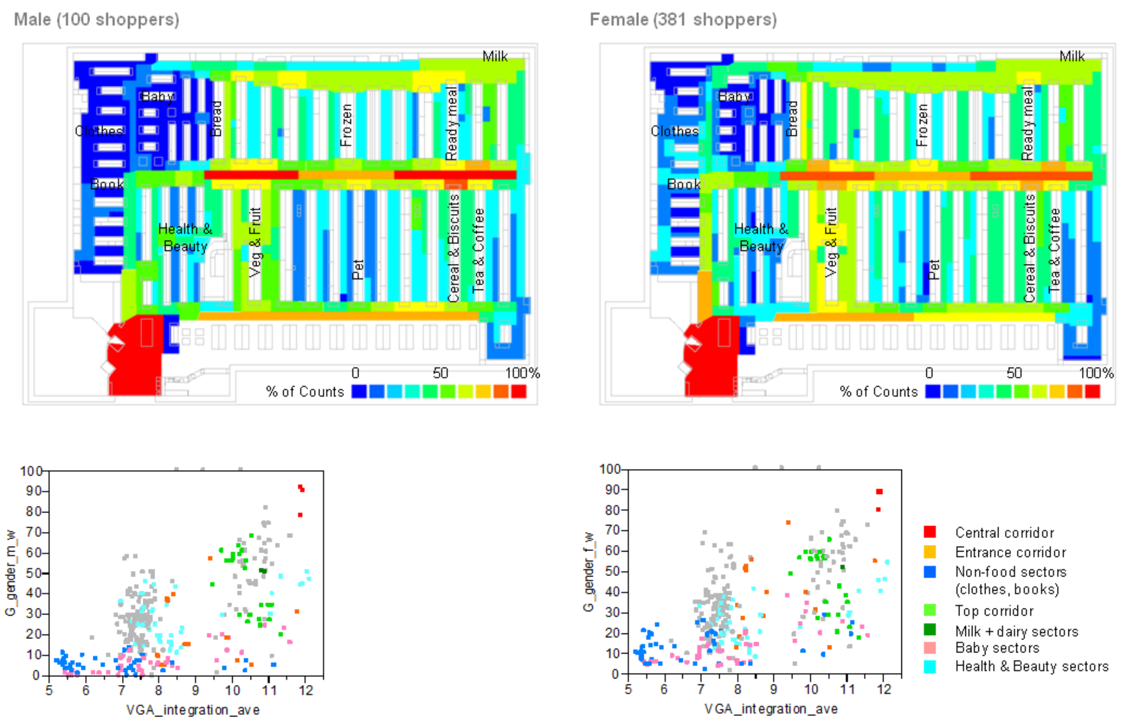

Figure 3
Example of visual and statistical analysis of movement levels in the store according to gender. The comparison between male and female shoppers doesn't show significant differences.

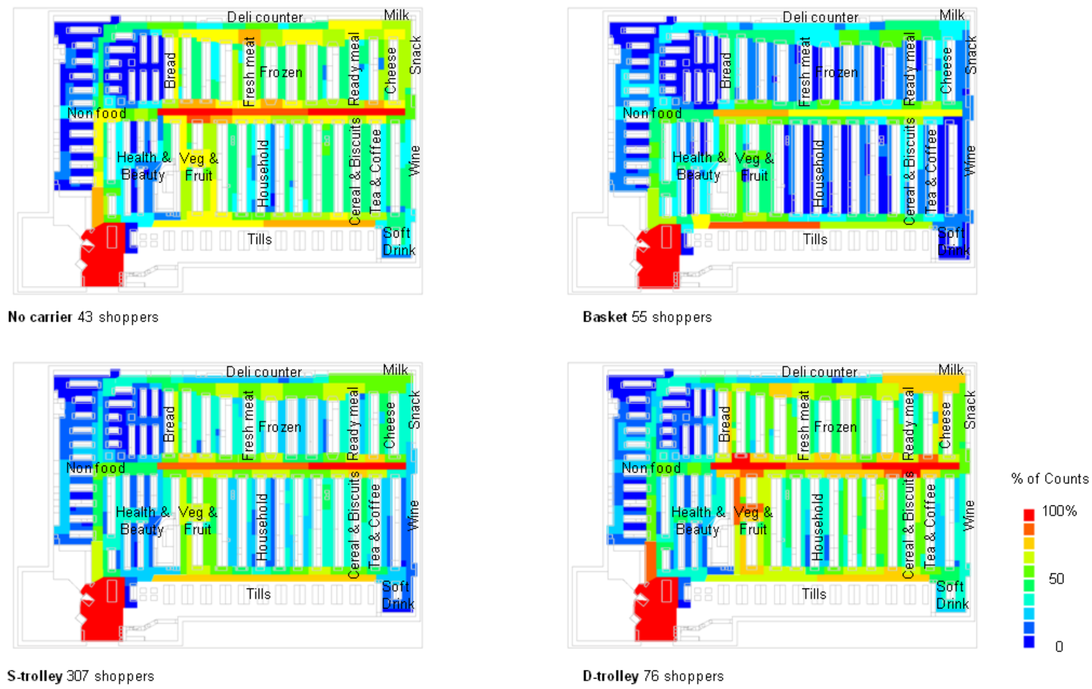

Figure 4
Comparison of movement levels in the store according to carrier type used

We then used the classification of shoppers according to each variable to produce filtered maps of the movement levels in the store, which are visually and statistically compared to identify any population differences within each variable. For example, we map the movement levels of male and female shoppers (Figure 3) to find that there is no significant difference in the use of the store, as opposed to the comparison between different carrier types: no carrier, basket, shallow trolley or deep trolley (Figure 4).

On the whole, few groups demonstrated significantly distinctive movement patterns or correlations. This suggests the complexity of shopper behaviour, which can hardly be explained by a single variable. Some groupings of shopper profile and shopping behaviour showed significant differences in their movement patterns between groups. However, it was not the case for any of the shopper demographic groups. Atypical groups demonstrated different movement patterns compared with the whole population, but there was no correlation with the other groupings. This analysis provided a basic understanding of relationships between shoppers' attributes and movement levels in the store, hinting that there are some groups of shoppers that have generally a different behaviour. To what extent can we identify different movement patterns that are defined by the spatial configuration of the store? And are these patterns specific to certain shoppers?

5. Spatial movement patterns

Based on a previous study by Larson et al. (2005) we cluster the movement traces to identify similar strategies to move around the store in order to complete a shopping trip. The clustering is made according to their location at a specific percentage of the shopping trip using a k-means algorithm. Each trace was segmented in 100 parts and the X and Y coordinates of each segmentation point were used as attributes of the trace for clustering, thus normalising duration and length of the trace. For each cluster obtained, the "medoid" is identified, that is the most representative trace of the cluster closest to its centre.

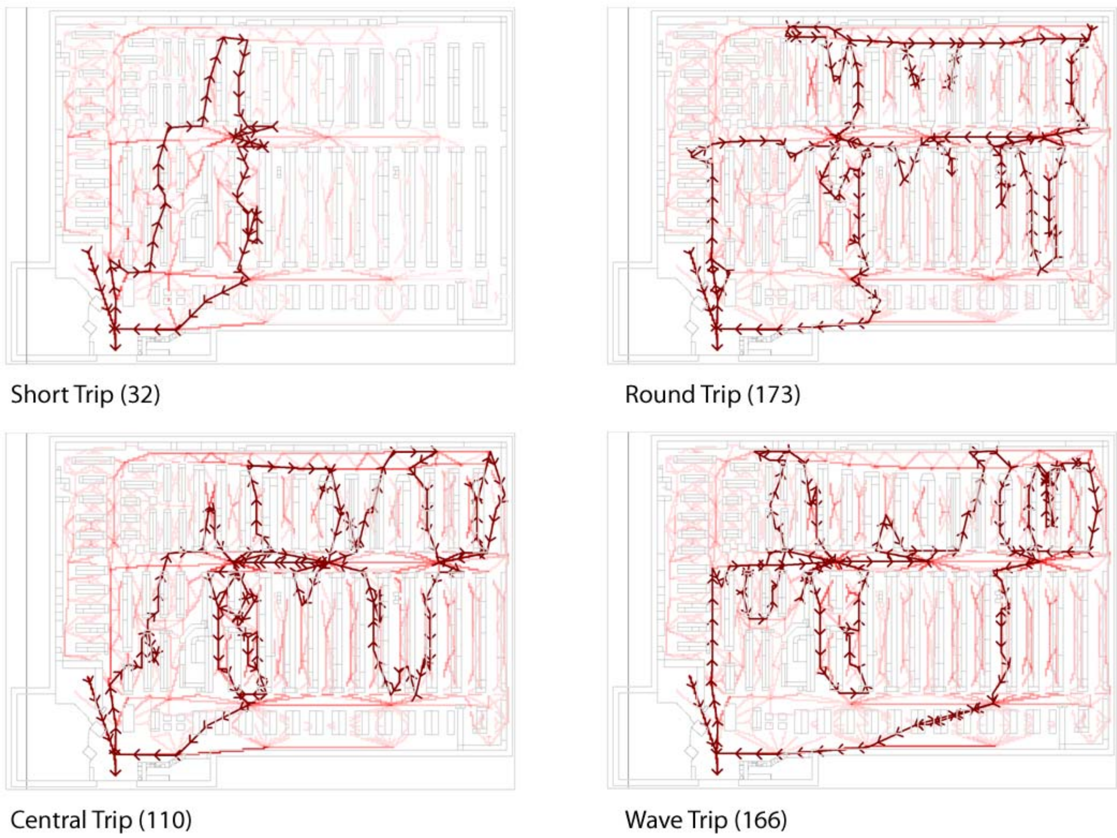

Figure 5
Plans representing each of the four movement pattern clusters. Each plan shows the combined level of movement from all traces in the cluster with varying intensity, overlaid with the medoid, where the arrows indicate the direction of movement.

With this method we were able to identify four distinctive movement patterns in the store, illustrated in Figure 5. Their description is derived from a detailed observation of the traces within each cluster using level of movement maps, the medoid and time series animations:

- Short Trip: 32 shoppers on a simple, short trip in and out of the store with few specific targets that can be located anywhere in the store, not necessarily visiting the most popular products.
- Round Trip: 173 shoppers moving up and along the top corridor and aisles, visiting the Vegetable, Fruit and Bread at the start; returning along the main corridor with various types of incursions into aisles; generally exit near fruit and vegetables.
- Central Trip: 110 shoppers progressing in and out of the store using the main corridor; with various types of incursions into the aisles, mainly visiting the top aisles first and the bottom ones when returning.
- Wave Trip: 166 shoppers in linear progression through the store along the main corridor, zigzagging through the aisles; most exiting near the far end of the store.

To understand the composition of each cluster, a detailed statistical profiling of shoppers within each cluster was created using charts that represent the "shopper DNA", based on different shopper profile attributes (Table 1).

Attributes	**0**	**1**	**2**	**3**	**4**	**5**	**6**
Gender	Male	Female					
Age group	-	18-24	25-34	35-44	45-54	55-64	65 and over
Group size	-	Alone	Two	Three	Four	Five	Six or more
Carrier type	None	Basket	Shallow trolley	Deep Trolley			
Frequency of visit	-	First Time	Regularly	Occasionally			
Shopping mission	-	Main	Top-up	Tonight	For Now	Non-food	
Shopping list	-	Yes	No				
Attitude to promotions	-	Always	Familiar	Familiar	Never		
Satisfaction	-	Very Satisfied	Satisfied	Neither	Dissatisfied	Very dissatisfied	
Shopping Duration	-	<10 min	<20 min	<30 min	<45 min	45 min or more	
Walking Speed	-	Slow	Medium	Fast			
Duration of Interactions	-	Short	Medium	Long			

Table 1
Shopper profile attributes used to build the shopper DNA of each cluster

By looking at the DNA profile of the shoppers in each of the four movement patterns (Figure 6), it becomes clear that apart from the "Short Trip" cluster, all other clusters fail to show significant differentiation in terms of shopper demographic and shopper profile. The "Short Trip" movement pattern represents shoppers using more baskets and no deep trolleys, none on a main shopping mission and with shorter shopping duration.

Although these movement strategies are clearly identified they are not representative of particular population groups. The choice of strategy turns out to be more an individual preference related to personal choice than to one's profile as shopper. Furthermore, these movement strategies are very much dependent on this specific shop layout, and as such they do not offer a more generic understanding of shopper behaviour types and how these lead to different use of space.

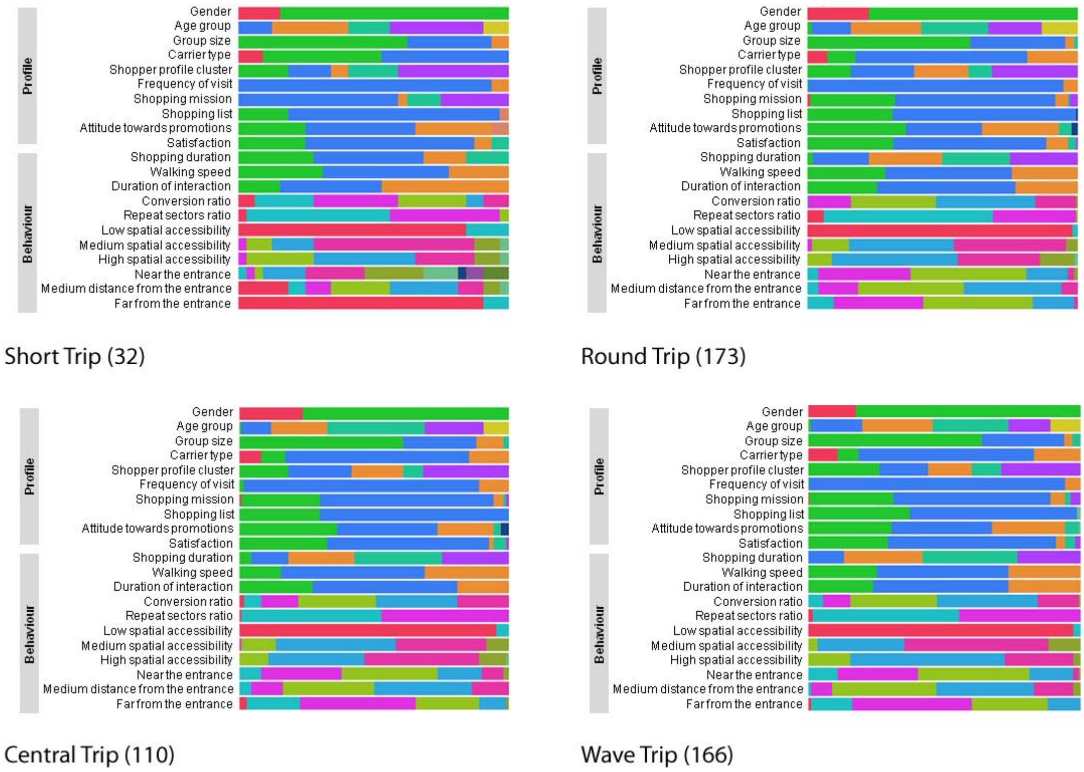

Figure 6
"Shopper DNA" profile of each movement pattern cluster. Each band represents the percent share of shoppers with a particular attribute according to the classification in Table 1 and Table 2. Apart from the "Short Trip" cluster, that has unique share of carrier types, shopping mission and shopping duration.

6. Spatial behaviour patterns

To enable us to identify the generic spatial behaviour types from the traces data set, we decide to take a different approach. Instead of using the exact trace route as the clustering attribute, we extract behavioural attributes from the traces data set and select a set of independent variables:

- Duration of shopping trip, in minutes
- Average walking speed, in metres per second
- Average duration of interactions with products, in seconds
- Conversion ratio, as the % of purchases from all interactions with products
- Repeat sectors, as the % of store sectors visited more than once
- Which areas of the store they visit, as the % of the trip in low, medium and high integration areas of the store
- How far into the store they go, as the % of the trip near, at medium distance and far from the entrance, in metric terms

Other behaviour attributes were considered, such as the total number of items purchased and the back tracking percentage, but these were found to be dependent variables and as such not considered for the clustering analysis. The selected behavioural attributes (Table 2) enable the description of how shoppers move and interact, instead of just where they go and when. This is similar to previous observational and descriptive studies of museum visitor types based on their spatial behaviour (Veron et al. 1983; Gabrielli 1999, Chittaro and Ienorutti 2004).

Attributes	0	1	2	3	4	5		
Shopping Duration	-	<10 min	<20 min	<30 min	<45 min	45 min or more		
Walking Speed	-	Slow	Medium	Fast				
Duration of interactions	-	Short	Medium	Long				
Percentages	0	10	20	30	40	50	60	70
Conversion Ratio								
Repeat Sectors								
Low Integration								
Med Integration								
High Integration								
Near Entrance								
Mid from Entrance								
Far from Entrance								

Table 2

Shopper behaviour attributes used to cluster and classify different types of shoppers. The variables with a percentage are represented by the range in the sample.

We run a k-means clustering analysis using the selected spatial behaviour attributes andcidentify five distinctive spatial behaviour patterns represented in Figure 7:

- The Specialist: 19 shoppers who focus on a few products, interacting with them for a long time. But those interactions are less likely to result in purchases. They are mainly on "top-up" or "non-food" mission.
- The Native: 161 shoppers make a long trip visiting only relevant aisles. Their interactions with products are likely to result in purchases. They are mainly on "main" or "top-up" mission.
- The Tourist: 101 fast moving shoppers, who prefer main corridors but don't go far from the entrance. They have a low conversion ratio looking more than buying. Some are on "non-food" missions.
- The Explorer: 67 shoppers making the longest trips, going everywhere more than once, slowly, with long interactions with the products and buying a lot. They cover all the aisles in the store, on a "main" shopping mission.
- The Raider: 113 fast shoppers, both in moving and making decisions, with clear preference for main corridors, going far into the store if necessary, to "top-up" or "food for tonight" missions. They have the highest ratio of male shoppers.

Because the attributes are not purely location based, the spatial patterns are less obvious than with the movement strategy clusters. However, the analysis of the "shopper DNA" charts using the shopper profile and shopper behaviour attributes (Figure 8) reveals clearer distinctions between the five clusters.

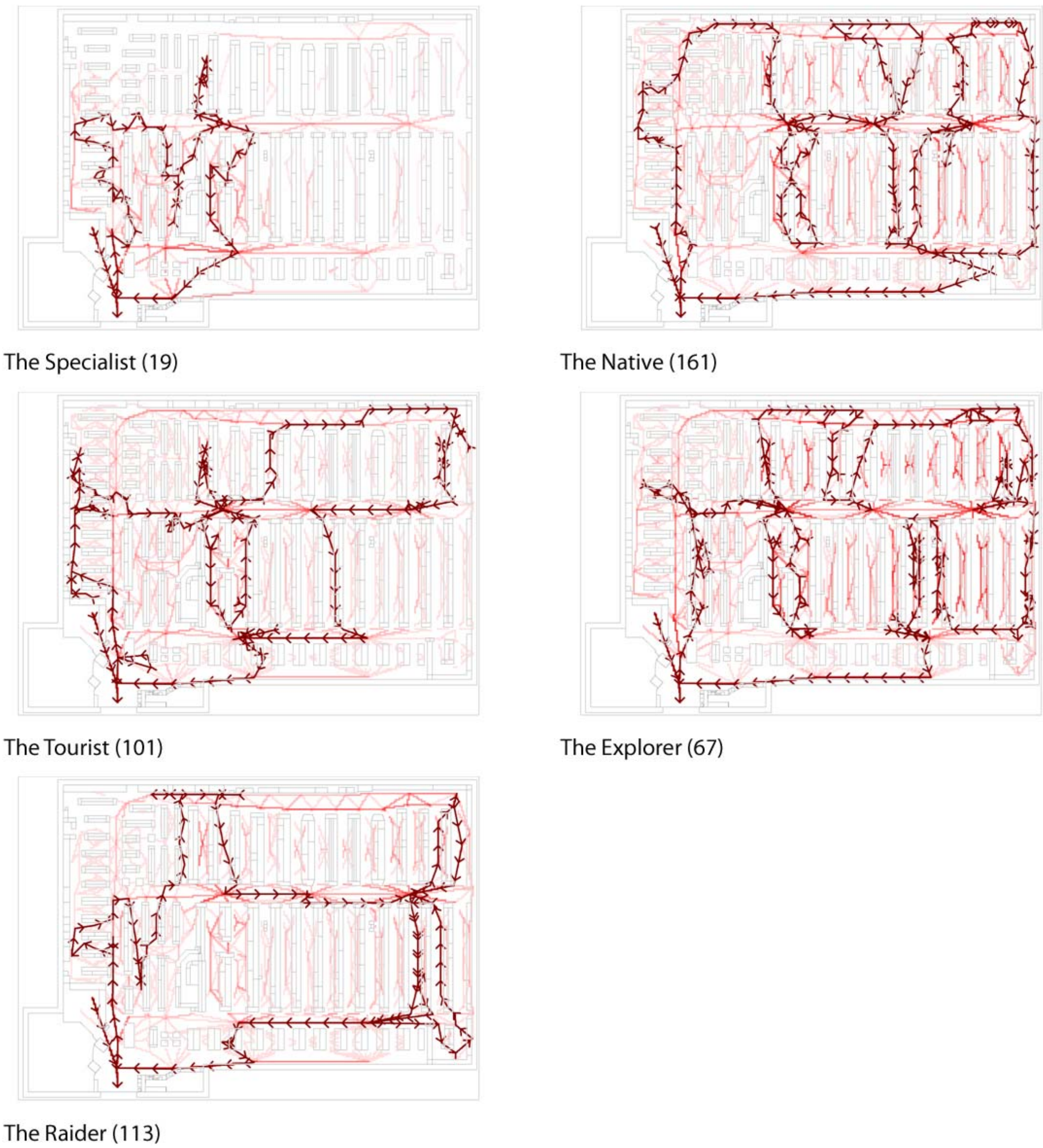

Figure 7
Plans representing each of the five spatial behaviour types, with the level of movement from all traces in the cluster displayed with varying intensity, overlaid by the "medoid", with the arrows indicating the direction of movement.

Below is a more detailed analysis of the constitution of each cluster in terms of both shopper profile and shopper behaviour.

The Specialist:

- 25% are male shoppers
- 58% use baskets rather and 30% shallow trolleys
- 42% are on "top-up" and 31% on "non-food" missions; 0% is on a "main" shop mission
- only 10% use a shopping list
- 20% of them are dissatisfied with the shopping experience
- 85% keep their shopping short to less than 20 minutes
- 47% interact with products for a long time
- only 5% of shoppers have conversion ratio above 30%
- 25% spend a part of the trip in the most segregated areas
- only 30% visit spaces which are far from the entrance
- 30% remain near the entrance for more than 80% of their trip

The Natives:

- 90% take a trolley
- 58% on “top-up” and 37% on “main” missions
- 34% use a shopping list
- 90% are satisfied with their shopping experience, of which 34% responded as “very satisfied”
- 85% take long and very long trips
- 52% have a conversion ratio above 40%
- Only 2.5% goes to segregated areas
- Spend 10% more of the trip in aisles than in corridors
- 40% of the trip in the middle of the store

The Tourist:

- 68% take trolleys
- 82% are on a top-up mission
- 33% have a shopping list
- Only 28% are very satisfied
- 35% are from the mature profile group
- 80% short and medium trips
- 36% are fast walking
- 80% of people with less than 40% conversion ratio
- Only 5% goes to segregated areas.
- 70% are more than 40% of the trip near the entrance
- 82% venture up to 20% of the trip far from the entrance and the remaining 18% don’t even go there.

The Explorer:

- 62% are female shopping alone
- 87% take a trolley
- 82% are on a main mission
- 43% have a shopping list, the highest share of all clusters
- 100% are very long trips
- 0% fast walking individuals
- 32% have long interactions with products
- 75% have more than 40% conversion ratio
- More distance covered in aisles than in corridors
- 56% does less than 30% of the trip near the entrance

The Raider:

- 33% share of male shoppers is the biggest
- 71% on “top-up” missions with 20% on “main” and “food for tonight” missions
- 75% do short and medium trips
- 100% walking at medium or fast speed
- 40% has short interactions
- 52% have a conversion ratio above 40%
- 0% visits to segregated areas
- 75% stay more than 50% of the trip in main corridors
- More time spent far than at medium and near distance from the entrance.

Overall, spatial behaviour patterns are easier to quantify and describe because they present unique identifiers in terms of shopper behaviour. But they also represent particular shopper profiles in terms of type of carrier, shopping mission and, to a lesser extent, gender. Furthermore, these behaviours are independent of store layout which can enable the use of these types to study different stores or different store layouts.

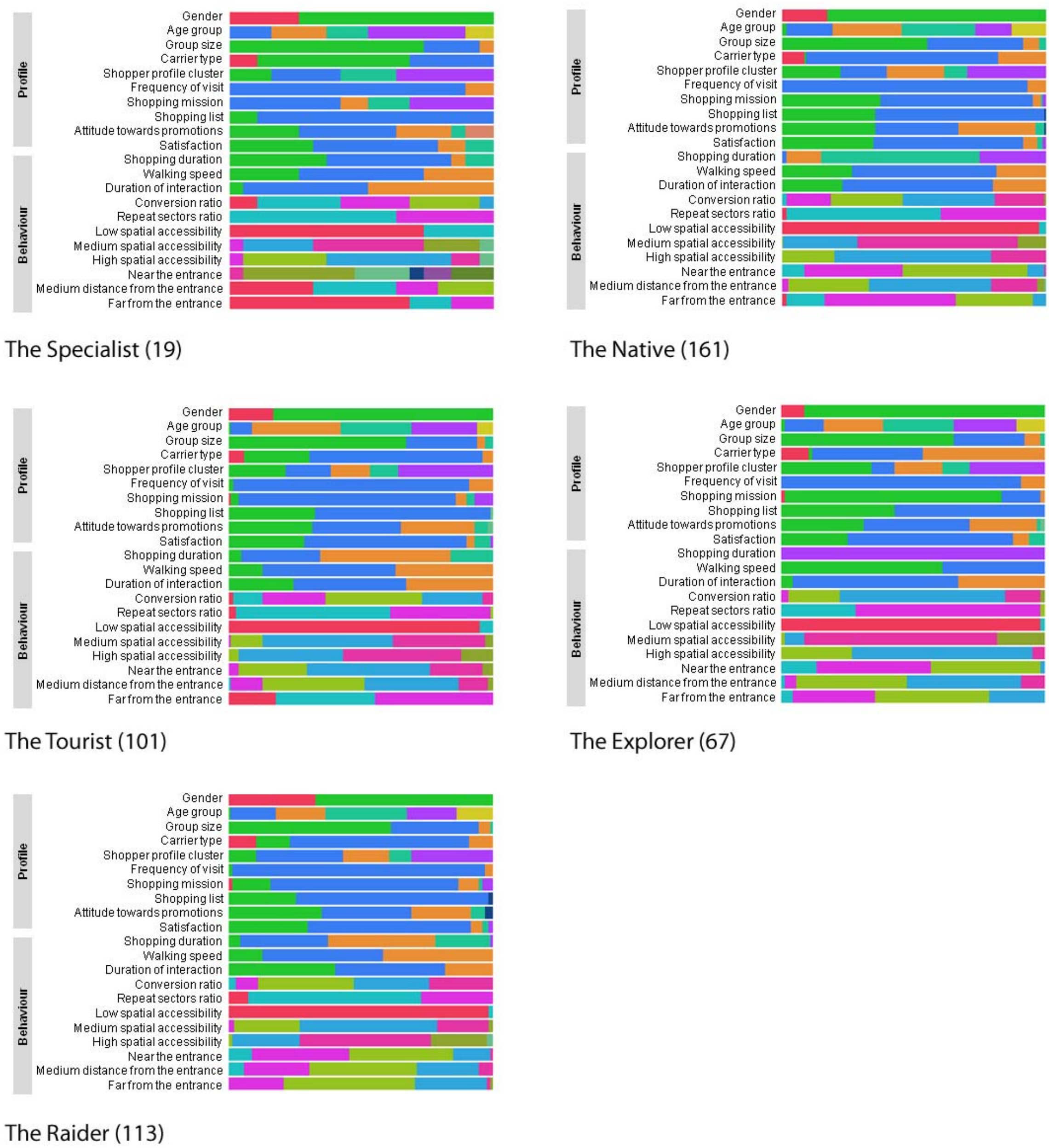

Figure 8

"Shopper DNA" profile of each spatial behaviour cluster

7. Conclusion

Our study provides an in-depth understanding of shopper behaviour in relation to a particular store layout. We envisage further validation of the findings by conducting comparable studies to cover a wider sample of store layouts, a wider variety of shopper profiles, changes of product location in the same store(s), and data on more first-time visitors. However, we believe that it this study also offers a more general set of tools for investigating the similarity and differences between individual trace data in spatial environments, and it provides good evidence that in the specific case of supermarkets, where goods are very evenly distributed in a regular and even grid, certain cases the homogeneity assumption is ill-founded. It seems clear that where the allocation of attractors to space is neither random nor correlated to spatial structure, different groups of people on different kinds of mission will adopt distinctive spatial strategies in their search for goods. In this case categories of goods are located in specific aisles in the supermarket in an organised and coherent way, but one that bears little relationship to the spatial layout of the store, which in any case is such a simple grid that there is little in the way of configurational information to differentiate one aisle from another.

References

Chittaro, Luca, and Lucio Ieronutti. 2004. A visual tool for tracing users' behavior in Virtual Environments. In *Proceedings of the working conference on Advanced visual interfaces*, 40-47. Gallipoli, Italy: ACM. doi:10.1145/989863.989868.

Gabrielli, F, P Marti, and L Petroni. 1999. The environment as interface. In *Proceedings of the i3AC'99*, 3:44-47.

Hillier, B., A. Penn, J. Hanson, T. Grajewski, and J. Xu. 1993. Natural movement: or, configuration and attraction in urban pedestrian movement . *Environment and Planning B: Planning and Design* 20: 29–66.

Hui, Sam, Peter Fader, and Eric Bradlow. 2008. The Traveling Salesman Goes Shopping: The Systematic Deviations of Grocery Paths from TSP-Optimality. Working Paper Series. *Social Science Research Network*. January 1.

Koch, Daniel. 2005. Parallel Spatial Scales - discerning cognitive levels of space. In *Proceedings of the 5th International Space Syntax Symposium*. Delft, the Netherlands: TU Delft, June.

Larson, Jeffrey S., Eric T. Bradlow, and Peter S. Fader. 2005. An exploratory look at supermarket shopping paths. *International Journal of Research in Marketing* 22, no. 4 (December): 395-414. doi:10.1016/j.ijresmar.2005.09.005.

Marcus, Lars. 2005. Plot Syntax - a configurational approach to urban diversity. In *Proceedings of the 5th International Space Syntax Symposium*. Delft, the Netherlands: TU Delft, June.

Markhede, Henrik, and Daniel Koch. 2007. POSITIONING ANALYSIS:
social structures in configurative modelling. In *Proceedings of the 6th International Space Syntax Symposium*. Istanbul, Turkey: Istanbul Technical University.

Miller, Daniel, Peter Jackson, and Michael Rowlands. 1998. *Shopping, Place and Identity*. Routledge.

Penn, A, and A Turner. 2004. Movement-generated land-use agglomeration: simulation experiments on the drivers of fine-scale land-use patterning. *URBAN DESIGN International* 9, no. 2: 81-96.

Peponis, J, E Hadjinikolaou, C Livieratos, and DA Fatouros. 1989. The spatial core of urban culture. *Ekistics* 56, no. 334/5: 43-55.

Peponis, J., R. Conroy Dalton, J. Wineman, and N. Dalton. 2003. Path, theme and narrative in open plan exhibition settings. In *4th International Space Syntax Symposium*. London, UK: UCL, June 17.

Peponis, John, Ruth Conroy Dalton, Jean Wineman, and Nick Dalton. 2004. Measuring the effects of layout upon visitors' spatial behaviors in open plan exhibition settings. *Environment and Planning B: Planning and Design* 31, no. 3: 453 – 473. doi:10.1068/b3041.

Read, S. 1999. Space syntax and the Dutch city. *Environment and Planning B: Planning and Design* 26, no. 2: 251 – 264. doi:10.1068/b260251.

Sailer, K. 2007. MOVEMENT IN WORKPLACE ENVIRONMENTS: configurational or programmed? In *Proceedings of the 6th International Space Syntax Symposium*, 068. Istanbul, Turkey: Istanbul Technical University.

Ståhle, Alexander, Lars Marcus, and Anders Karlström. 2005. Place Syntax- Geographic Accessibility with Axial Lines in GIS. In *Proceedings of the 5th International Space Syntax Symposium*. Delft, the Netherlands: TU Delft, June. doi:10.1.1.109.678.

Ståhle, A., L. Marcus, and A. Karlstrom. 2007. Place Syntax Tool—: GIS Software for Analysing Geographic Accessibility with Axial Lines. In *New Developments in Space Syntax Software*, ed. Alasdair Turner, 35–42. Istanbul, Turkey: Istanbul Technical University.

Stutz, Chris, Jorge Gil, Eva Friedrich, and Corey Klaasmeyer. Confeego. London, UK: Space Syntax Limited. http://www.spacesyntax.org/software/newtools.asp.

Turner, A. 2003. Analysing the visual dynamics of spatial morphology. *Environment and Planning B: Planning and Design* 30: 657-676.

Turner, Alasdair. UCL Depthmap: Spatial Network Analysis Software. London, UK: UCL. http://www.vr.ucl.ac.uk/depthmap/.

Veron, E, and M Levasseur. 1983. *Ethnographie de l'exposition bibliotheque publique d'information*. bibliotheque publique d'information. Paris: Centre Georges Pompidou.