Environment and Planning A, 1988, volume 20, pages 891-906

Space-time diaries and travel characteristics for different levels of respondent aggregation

D G Janelle, M F Goodchild
Department of Geography, University of Western Ontario, London, Ontario, N6A 5C2, Canada
B Klinkenberg
Department of Geography, University of British Columbia, 1984 West Mall, Vancouver, V6T 1W5 Canada
Received 27 January 1987; in revised form 10 August 1987

Abstract. Significant progress has been made in the analysis of space-time diary data. Drawing on the flexibility that such data provide, in this study the authors group respondents at five different levels of aggregation, and compare them according to their mean and standard deviation values for selected measures of travel behaviour. The measures, derived from the time-geography model, relate to the range and speed of daily movement and to the duration of activities. Wide variation in values were observed among subpopulations and role groups at each level of aggregation and, in general, these increased for higher levels of disaggregation. Graphic plots of the mean and standard deviation values permit evaluations of the effects of aggregation and provide a basis for identification of relationships between respondents' socio-demographic characteristics and their travel behaviour.

Space-time diaries provide a basis for detailed description and analysis of individual travel behaviour (Anderson, 1971). They record the locations of individuals continuously throughout the day, and it is possible to use them for the reconstruction of the timing, sequence, duration, and frequency of activities and trips. Diaries usually contain elaborate classifications of activity and are accompanied with surveys on the sociodemographic characteristics of respondents, and thus they provide an opportunity to consider travel behaviour in the broader context of the individual's patterns of activity. Significant steps in these directions are represented in the recent literature. Burnett and Hanson (1979) and Pas (1983) have used diary data to describe the complexity of travel linkages and trip chains, and Dangschat et al (1982), Hanson (1982), Hanson and Hanson (1980, 1981), Janelle and Goodchild (1983), and Pas (1984) have attempted to characterize the social group variations and individual traits which are associated with travel patterns and trip indicators.

Notwithstanding the problems in the handling of space-time diary data, significant technical difficulties have been resolved. These include (a) the ability to reconstruct the space-time paths for individuals (Burnett and Hanson, 1979), allowing for the direct inclusion of empirical evidence on the constraints specified in Hägerstrand's time-geography model (Burns, 1979; Hägerstrand, 1974; Lenntorp, 1978), and (b) the designation of subpopulations, which is based on constructs of role and life cycle that have relevance for the interpretation of patterns of travel behaviour (Brail and Chapin, 1973; Burnett and Hanson, 1979; Clark et al, 1980; Harvey et al, 1977; Horton and Hultquist, 1971; Janelle and Goodchild, 1983; Kutter, 1973; Oppenheim, 1975).

Both of these accomplishments are drawn upon in this paper. Here, the general objective is to measure attributes for the space-time paths of different subpopulations. The attributes include indicators of the geographic range and speed of daily movements, as well as measures of activity duration. Many of the measures are keyed to important locations, such as the home, workplace, city centre, and shopping locations—some of the major anchor points of daily activity.

In the paper, our specific objective is to document and to interpret variations in the calculated values for different measures based on the level of aggregation of respondents. Drawing on the considerable flexibility that diary data provide for grouping respondents by a large number of characteristics, the mean and standard deviation values, for selected travel indicators, will be compared at different levels of a priori aggregation, and according to groupings defined by a multivariate classification routine.

The data

This study is based on the analysis of more than 1200 geographically coded single-day space-time diaries for a random sample of respondents in Halifax, Canada. Detailed descriptions of the survey design and documentation on its reliability are provided by Elliott et al (1976). The original survey included 2141 people, but this study is based on the subset who lived and worked within the confines of the municipal boundaries of Halifax and its twin city, Dartmouth, and who maintained diaries on weekdays as opposed to weekends. The activities of each individual are described in line with the detailed classification of the International Time-Budget Study (Szalai, 1972), all locations occupied during the diary day are geographically-coded to the nearest 0.1 kilometre, and a detailed sociodemographic survey of each respondent was completed.

There are three forms of aggregation which is relevant to the analysis of space-time diary data; events might be aggregated by time interval, or spatially, or by the characteristics of individuals. In this investigation, all travel indicators are aggregated at the daily level (for example, number of trips per day), and space is represented by distance measures (for example, kilometres per trip or maximum distance from the city centre on diary day). However, in keeping with the objective in this paper, the grouping of respondents will range from simple dichotomous categories (for example, male or female) to combinatorial groupings of up to six dichotomous variables, to groupings based on a cluster analysis of fourteen criterion variables. The rationale for the selection of variables and for the combinations of variables will be presented after a brief description of the travel indicators.

Travel indicators

Presented in table 1 is a selection of seventeen measures, divided into constructs based on the geographic range of daily trip making, or on the capability for movement within the urban environment, or on the physical accessibility of respondents to key locations, and on activities and durations of travel episodes. The significance of these four constructs was confirmed by means of principal-axis factor analysis, as reported in previous research (Janelle and Goodchild, 1983). Shown in table 1 are the means and standard deviations for values of the measures which are based on the highest possible level of aggregation—all individuals are treated as equivalent members of a single grouping. These values will be treated as the base for comparison of the same measures derived at lower levels of aggregation.

In general, each of the specific indicators may be seen as representative of the level of constraint on the activity behaviour of respondents, reflecting the dominant coupling and capability constraints embraced in Hägerstrand's time-geography model. For example, geographic range relates to the spatial extent of one's space-time path for travel events; the measures of physical accessibility focus on the importance of the home base and work stations in the constraint of movements according to the biological and social requirements of home life and to the need for participation in the work economy; and the numbers and durations of activities give some indication of the complexity of daily programs. Although the time-geography model has been used in a diagnostic sense, these derivative measures provide neither simple arithmetic indicators of constraint levels on behaviour nor unambiguous surrogates for qualify-of-life interpretations. For example, whereas a high geographic range may reflect a varied exposure to geographically distributed opportunities and good access to transport technology, it may also be associated with environments which have poor provision of many services, thus necessitating lengthy trips or the careful selection of residential and other locations in order to hold travel expenditures in check. The resolution of this kind of ambiguity requires detailed information on the circumstances and preferences of specific individuals. Alternatively, some clarification may be derived from a systematic focus on these measures for the disaggregated groupings of the sample respondents.

Table 1. Averages for selected travel indicators, calculated from space-time diaries, Halifax, Canada.

Travel indicators[a]	Mean	Standard deviation	Sample size[b]
1 *Geographic range* (km)			
Maximum distance from home on diary day	4.1	4.3	803
Maximum distance from city centre on diary day	5.3	4.8	1181
Maximum distance from primary workplace on diary day	4.5	4.3	786
Total distance of all travel episodes on diary day	9.6	13.5	1207
Mean distance per trip	1.8	2.2	1112
2 *Transportation capability* (km h^{-1})			
For all trips regardless of mode	8.5	9.3	1091
For work trips by auto	17.9	10.4	573
For work trips by bus	8.3	7.8	321
3 *Physical accessibility*			
Distance from home to work (km)	3.3	2.9	787
Distance from home to usual shopping location (km)	3.0	2.5	1072
Commuting time by private auto (minutes)	13.3	7.8	601
Commuting time by bus (minutes)	29.9	17.8	323
4 *Activities and durations*			
Number of separate activities listed in diaries	29.2	9.4	1207
Duration of all travel episodes (minutes)	68.1	53.5	1207
Duration per travel episode (minutes)	13.5	9.1	1092
Discretionary hours per day	4.6	2.7	1185
Total number of trips per day	5.5	3.7	1207

[a] Calculated with data from Halifax space-time budgets (Elliott et al, 1976). All distance and speed measures are based on straight-line paths between activity sites.

[b] Sample sizes will vary depending on the number of missing values and on the number of respondents who qualify for a given measure. For example, distances from the workplace are only relevant for employed respondents.

The designating of subpopulations and role groups

Much of the literature on travel behaviour is based on single-variable subpopulations. In table 2, the study sample is divided by six dichotomous variables into twelve of these subpopulations. The importance of these groupings has been confirmed by behavioural researchers such as Brail and Chapin (1973), Hanson and Hanson (1981), and Harvey et al (1977). But, in addition, research in time-geography has confirmed their significance, in terms of coupling and capability constraints (Lindsay, 1979; Palm and Pred, 1974). Thus, employment status influences the selection of activities through income constraints; marriage, childcare responsibility, jobs, and needs for home maintenance define obligations which require the allocation of time and the occupance of specific locations; and the availability of an automobile may circumscribe the geographical field and timing of one's activities. It is expected that the extent of these constraints will be revealed by the transport indicators listed in table 1.

Aside from the single-variable subpopulations, also defined in table 2 are four-variable and six-variable role groups. Even though these fall short of describing the full complexity of real human beings, they are assumed to provide greater behavioural relevance than single-characteristic subpopulations. That is, the category FSNERA (a single, employed female with no children, who rents her accommodation and has access to an automobile) captures the essence of the person more completely than would any one of the six component traits if they were treated separately.

Disaggregations to the four-variable level are based on a suggestion from Clark et al (1980) for the combination of attributes of sex, marital status, childcare responsibility, and employment into a typology of role groups. Although Dangschat et al (1982) use empirical findings to raise doubts about the validity of the 'groups of homogeneous behaviour' approach, (but compare their results with Hanson and Huff, 1986), the literature, in general, offers evidence for the importance of these particular attributes (for example, Hanson and Hanson, 1980; Harvey et al, 1977). In the a priori selection of attributes for inclusion in the combinatorial classification process, one can draw upon empirical findings and from conceptual insights of the relationships between individuals and their social settings. For example, the view of Cullen and Godson (1975), that routine daily behaviour is largely beyond individual control except in the context of important life-decisions, provides a basis for grouping

Table 2. Sample sizes for subpopulations and role groups.

Single-variable subpopulations	Size	Role groups			
		four variables	size	six variables	size
Female (F)	675	FWNE	71	FWNEOA	23
Male (M)	532			FWNERA	40
Married (W)	886	FWNU	66	FWNUOA	28
Not married (S)	312	FWCE	104	FWCEOA	55
				FWCERA	28
No children in household (N)	438	FWCU	236	FWCUOA	133
Children in household (C)	730			FWCURA	63
Employed (E)	791	FSNE	105	FSNERA	31
Unemployed (U)	416			FSNERP	48
Homeowners (O)	616	FSCE	33		
Renters (R)	569	MWNE	97	MWNEOA	40
Autos/trucks in household (A)	917			MWNERA	41
No autos/trucks (P)	227	MWCE	264	MWCEOA	134
				MWCERA	95
		MWCU	22		
		MSNE	57	MSNERA	22
		MSCE	26		
		MSCU	23		
		other categories[a]	55	other categories[a]	363
		missing cases	48	missing cases	63
Total number of respondents	1207		1207		1207

[a] Include groupings with fewer than twenty-two respondents.

respondents according to such decisions; to their marriage, job, childcare, and homeownership characteristics. In addition, the six-variable combinatorial groupings shown in table 2 include access to the use of an automobile by respondents.

The concatenation of four and six dichotomous variables produces sixteen and sixty-four possible groupings, respectively; however, the analysis is based on only those twelve four-variable and fourteen six-variable role groups which have at least twenty-two members. For each of these role groups, and for each of the twelve subpopulations, we calculated the means and standard deviations for four selected travel indicators. Each indicator is representative of one of the constructs listed in table 1: average maximum kilometres from home, for geographic range; average speed for all trips, for transportation capability; average total minutes for all trips, for activities and durations; and average kilometres to usual shopping location, for physical accessibility.

Significance levels were calculated for each travel indicator. For the single-variable analyses, each pair of subpopulations was compared by use of a two-sample t-test. For analyses in which the population was broken into more than two groups, two-sample t-tests were used to compare each subpopulation with the remainder. For example, the mean distance travelled to a usual shopping location, for the four-variable role group FWNE (female, married, no children, employed), was compared to the mean distance for all other individuals. In both forms of the test, the null hypothesis was that both the subpopulation and the remainder were drawn randomly and independently from the same, hypothetical, normally-distributed population.

The two-sample t-test is subject to a number of assumptions, and the resulting significance levels should not be taken as definitive. It is possible for subpopulations to have similar means but to differ in variances; large differences in variance were observed in many of the analyses. It is likely also that many subpopulations show non-normal distributions for some of the travel indicators.

The role groups based on one, four, and six simple dichotomies provide a straightforward way to relate variation in travel behaviour to individual characteristics of the sample. Although the individuals in each group are identical in having the characteristics associated with the group, they are not necessarily homogeneous with regard to other characteristics, particularly if these latter traits are statistically independent of the former. For comparison, the sample was analyzed by use of the fourteen standardized and equally weighted criterion variables which are listed in table 3. By use of cluster analysis which incorporated Ward's algorithm (Ward, 1963; Wishart, 1982) ten homogeneous groups were produced. Because of a limit to the number of cases that could be analyzed, a random sample of 995 individuals was clustered; the remaining 212 cases were then assigned to the ten groups by discriminant analysis.

The means of each group for the fourteen criterion variables are shown in table 3; the means of an additional ten variables are shown to aid in the general description of each group. The ten groups were then used as a fourth aggregation of the sample, for comparison with the three sets of role groupings described previously. Although the ten clusters are arguably more homogeneous than any of the previous three sets, the results are nevertheless not as easy to interpret because cluster membership is based on continuous variables and on generalized concepts of distance rather than on simple dichotomies.

In addition to a priori grouping and cluster analysis, the sample could have been partitioned by the use of segmentation techniques such as AID (Automatic Interaction Detection) (Hensher and McLeod, 1977); by a search for that combination of individual traits which partition the sample into groups that are maximally homogeneous with respect to a dependent variable, in this case some aspect of

Table 3. Average values of criterion and descriptive variables for groups defined by cluster analysis.

	Group[a]										Total sample
	1 (34)	2 (65)	3 (152)	4 (110)	5 (134)	6 (166)	7 (135)	8 (163	9 (96)	10 (152)	(1207)
Criterion variables[b]											
Number of children <18 years old	1.2	2.3	0.5	0.8	0.7	1.1	3.0	2.8	0.9	0.8	1.4
Age (years)	21.6	28.3	29.6	31.6	33.6	34.2	37.2	40.2	50.2	54.4	37.6
Highest grade in school	11.8	10.3	11.0	10.8	10.3	11.1	10.0	10.9	9.5	9.6	10.5
Number of years at university	2.1	1.1	0.3	0.8	5.1	2.5	0.2	1.5	0.1	0.4	1.4
Monthly spending on rent and utilities ($)	168	189	149	160	147	252	209	217	155	152	184
Household income ($)	8750	7188	6769	5775	6776	14102	9064	12782	9572	8468	9359
Rooms in dwelling	5.4	4.9	3.5	3.9	3.9	6.3	5.9	6.6	5.2	5.9	5.2
Number of cars/trucks per household	1.0	0.9	1.2	0.7	0.7	1.5	1.1	1.4	1.1	1.1	1.1
Number of other vehicles per household	0.7	0.7	0.1	0.2	0.3	0.4	1.5	1.9	0.3	0.3	0.7
Diary-day activities											
Work for income	0.7	0.5	6.4	0.3	4.8	4.4	0.3	6.4	7.1	0.4	3.5
Attend classes	4.3	0.0	0.03	0.0	0.03	0.01	0.0	0.11	0.0	0.0	0.13
Housework	0.3	4.1	0.9	2.2	1.4	1.5	5.3	0.9	0.8	3.4	2.1
With children	0.01	3.3	0.1	0.1	0.1	0.3	0.8	0.2	0.04	0.1	0.4
Sleeping	7.4	7.7	7.3	9.8	7.0	7.6	7.8	7.0	7.0	7.8	7.6
Additional descriptive variables											
Dwelling with >5 rooms (%)	38.2	29.2	5.9	12.7	18.2	61.8	58.6	69.8	41.7	64.2	42.4
Monthly spending on rent and utilities ⩾$200 (%)	28.0	29.5	14.7	19.0	15.0	80.7	56.3	60.9	25.0	26.2	38.2
Household income ⩾$12000 (%)	32.4	13.8	11.2	8.2	14.9	62.0	31.1	54.6	20.8	27.6	30.0
⩾1 year spent at university (%)	64.7	23.1	16.4	21.8	94.8	56.0	12.6	37.4	4.2	10.5	33.5
Number of females (%)	29.4	92.3	56.6	62.7	54.5	44.6	91.9	33.1	21.9	68.4	55.9
Number of employed (%)	20.6	16.9	97.4	58.2	82.1	78.3	17.0	94.5	100.0	31.6	65.5
Hours worked per week (for employed)	24.8	32.0	39.8	36.8	39.7	40.1	26.7	42.8	39.9	34.7	39.1
Years in present job	3.8	2.5	3.9	5.1	5.0	6.9	3.8	9.2	11.6	14.2	7.3
Number presently married (%)	26.5	93.8	51.0	63.3	53.7	75.0	93.2	89.6	88.5	80.7	74.0
Households without autos/trucks (%)	20.6	23.4	40.9	36.1	31.5	4.9	14.1	3.7	5.3	17.6	19.1

[a] Sample size is given in parentheses.

[b] Variables used to define clusters by Ward's algorithm (Ward, 1963; Wishart, 1982).

travel behaviour. However, a priori grouping and clustering seem more consistent with the broad objectives of the study.

In contrast to the subpopulations and role groups, which may be seen as more arbitrary categories, those based on cluster analysis may be viewed as statistically more 'natural' divisions of the respondents. The criterion variables were selected to reinforce this interpretation by the inclusion of representative measures of life cycle (age and number of children), socioeconomic status (education and income), housing (rooms per dwelling and expenditures), mobility (autos and other vehicles), and activities (hours devoted per day to selected kinds of episodes). All of these variables are judged to have influence on one's travel behaviour, and to represent different types and degrees of coupling and/or capability constraints.

To give a comparison of the different methods of dividing the sample, shown in table 4 are the within-group mean sum of squares for each of the four travel variables and for each of the methods used. In each case, the within-group sum of squares was divided by its degrees of freedom (sample size − number of groups) to give a measure which is less dependent on sample size. This statistic is expected to equal the variance of the entire sample under a null hypothesis which states that each group is sampled independently from the same parent distribution. Thus, relatively small values would be interpreted as indicating that the particular method of division had created groups which were relatively homogeneous on the particular variable of travel. In general, the values are expected to decrease as the number of groups increases, other things being equal.

Several trends are apparent from table 4. The cluster groupings give relatively low values on most variables, which indicates that, although these groups were constructed on the basis of a set of socioeconomic variables, they are relatively homogeneous on the travel variables as well. The relatively large values for the six-variable role groupings are the result of a calculation based on the fourteen well-populated groups rather than the full sixty-four. The sixteen four-variable role groups are substantially more homogeneous for the mean time in travel, than are the ten clusters, but otherwise values are similar, which indicates that the ten groups formed by clustering on the socioeconomic variables are as effective in reducing variation in travel behaviour as are the sixteen groups formed on explicit individual traits. Last, there are cases, such as the rent/own cleavage and the mean distance to a shopping location, in which a simple binary classification on one variable is as effective at reducing variance for a particular travel variable as is clustering or either method of role grouping.

Table 4. Within-group mean sums of squares for each travel variable and each method of sample aggregation.

Methods	Number of groups	Maximum distance from home	Speed for all trips	Total time for all trips	Distance to usual shopping location
Cluster	10	17.8	82.7	0.740	6.05
Four-variable role groups	16	18.3	82.6	0.715	6.08
Six-variable role groups	14	19.5	94.7	0.763	6.90
Male/female	2	18.0	84.7	0.756	6.14
Married/unmarried	2	18.1	84.8	0.773	6.06
Children/no children	2	18.4	85.5	0.792	6.17
Employed/unemployed	2	18.1	83.9	0.754	6.13
Rent/own	2	18.1	85.7	0.803	6.03
Auto/no auto	2	17.5	86.5	0.805	6.07

At this stage, a common pool of representative individuals in Halifax has been disaggregated according to four different classification schemes. The groups identified in each classification share personal traits with those of the other classifications. For example, the male (M) subpopulation is represented in six of the four-variable role groups and in five of the six-variable role groups, and is dominant in four of the cluster analysis groupings (1, 6, 8, and 9). Central to this research is the question of how these disaggregations of males and other basic subpopulations contribute to variations in their travel behaviour.

Analysis

Listed in tables 5, 6, 7, and 8 are the respective means, standard deviations, and significance levels for the subpopulations, four-variable role groups, six-variable role groups, and cluster classifications. Although the distance from a usual shopping location can be treated as a general indicator of accessibility to services from one's dwelling place, the mean values for travel range, speed, and trip duration should be interpreted with care. Whereas they may be seen to represent the level of constraint faced by members of a category, the 'quality-of-life' significance of high versus low mean values is not immediately apparent from the evidence at hand. Thus, from table 5, although the average male travels further, at higher speeds, and for longer periods than the average female, it should not be inferred that one group is, in any sense, 'better off' than the other. Rather, it is likely that individuals trade-off different modes of travel with alternative locations, depending on specific circumstances with respect to a large number of characteristics, needs, and preferences. The high standard deviations shown in table 5 for single-variable subpopulations reflect the levels of heterogeneity among the respondents within groups. At higher levels of disaggregation (tables 6, 7, and 8) the range of mean and standard deviation values tends to increase, adding to the ambiguity of interpretations. To facilitate the interpretation, the values for all respondent groupings are plotted on a single graph for each of the measures (figures 1-4). This procedure permits an assessment of the effects of aggregation, and also provides a basis for observing the relationships of different combinations of personal characteristics with measures of travel behaviour.

Table 5. Selected average daily travel indicators for single-variable subpopulations[a].

Subpopulation	Maximum distance from home (km)	Speed for all trips (km h^{-1})	Total time for all trips (minutes)	Distance to usual shopping location (km)
Female	3.6* (3.5)	7.5* (9.3)	57.2* (47.7)	3.0 (2.5)
Male	4.6* (4.7)	9.6* (9.1)	81.5* (57.3)	2.9 (2.4)
Single	3.5* (3.7)	6.7* (7.8)	83.6* (65.6)	2.4* (2.0)
Married	4.4* (4.5)	9.2* (9.7)	62.8* (47.4)	3.1* (2.6)
No children	3.8* (3.7)	8.0** (8.3)	71.3* (59.0)	2.8* (2.1)
Children at home	4.5* (4.7)	8.9** (9.8)	65.9* (49.7)	3.1* (2.7)
Employed	4.2* (4.2)	9.4* (9.5)	76.9* (54.2)	2.9** (2.5)
Unemployed	2.7* (4.8)	6.4* (8.3)	51.2* (47.8)	3.1** (2.5)
Renters	3.7* (4.1)	7.6* (8.4)	67.1 (56.6)	2.6* (2.3)
Homeowners	4.7* (4.5)	9.5* (10.0)	69.3 (51.0)	3.3* (2.6)
Autos	4.4* (4.3)	9.3* (9.7)	68.4 (54.5)	3.1* (2.6)
No autos	3.0* (3.4)	5.8* (7.5)	64.7 (51.0)	2.2* (2.3)

[a] Standard deviation is given in parentheses.

* Significant at 0.05 level. ** Significant at 0.01 level.

The results for the maximum distance that a person travels from home on a weekday illustrate very clearly how aggregate groupings may mask a wide scatter of internal variations. As seen in figure 1, the means for the disaggregate groupings

Table 6. Selected average daily travel indicators for four-variable role groups[a].

Role group[b]	Maximum distance from home (km)	Speed for all trips (km h^{-1})	Total time for all trips (minutes)	Distance to usual shopping location (km)
FWNE	4.3 (3.3)	10.6* (9.3)	65.1 (38.5)	2.7 (2.1)
FWNU		7.1** (10.7)	49.0* (51.6)	3.0 (2.3)
FWCE	3.7* (3.8)	9.0 (10.2)	65.2 (43.2)	3.2 (2.9)
FWCU	2.7 (4.6)	6.6* (8.5)	44.2* (41.9)	3.4* (2.7)
FSNE	3.2* (2.8)	6.3* (5.7)	73.7 (52.8)	2.7 (2.4)
FSCE	4.3 (4.6)	9.8 (14.9)	71.7 (46.6)	2.4** (1.9)
MWNE	4.3 (4.6)	10.5* (9.6)	75.8** (50.5)	3.0 (2.2)
MWCE	5.0* (5.0	11.0* (10.0)	77.6* (47.7)	2.9 (2.7)
MWCU	4.2 (6.9)	7.8 (10.6)	48.1* (35.9)	4.5* (2.7)
MSNE	3.8 (4.8)	6.8** (6.1)	102.1* (95.2)	2.6 (1.9)
MSCE	4.5 (2.8)	7.9 (6.5)	112.8* (65.2)	2.4** (2.1)
MSCU		6.2** (5.4)	104.5* (49.8)	1.6* (1.0)

[a] Standard deviation is given in parentheses. Omissions indicate missing data or small sample size which precluded valid calculations.

[b] F female, M male, W married, S single, C children at home, N no children, E employed, U unemployed.

* Significant at 0.05 level. ** Significant at 0.01 level.

Table 7. Selected average daily travel indicators for six-variable role groups[a].

Role group[b]	Maximum distance from home (km)	Speed for all trips (km h^{-1})	Total time for all trips (minutes)	Distance to usual shopping location (km)
FWNEOA	5.0 (2.8)	11.5 (7.7)	66.1 (32.1)	3.0 (2.4)
FWNERA	4.3 (3.6)	10.9 (10.5)	66.2 (41.4)	2.6** (2.0)
FWNUOA		11.0 (14.6)	56.6 (61.5)	3.1 (2.4)
FWCEOA	3.9 (3.6)	10.4 (11.4)	63.5 (42.0)	3.6** (3.4)
FWCERA	3.8 (5.1)	8.3 (10.3)	68.3 (54.5)	2.6 (2.3)
FWCUOA	3.8 (6.0	7.8* (8.7)	51.0* (45.5)	3.8* (2.9)
FWCURA		6.1* (9.5)	33.8* (34.9)	2.7** (2.0)
FSNERA	4.0 (3.2)	6.5* (5.5)	81.3** (46.0)	3.4 (3.6)
FSNERP	2.9* (2.7)	5.3* (5.4)	78.8** (63.8)	2.3* (1.6)
MWNEOA	5.0 (3.5)	11.8** (8.3)	76.7 (39.8)	3.5 (2.6)
MWNERA	4.0 (5.1)	10.9 (11.7)	75.1 (59.5)	2.6 (1.8)
MWCEOA	5.3* (5.4)	12.1* (11.9)	85.0* (49.0)	3.2 (2.6)
MWCERA	4.4 (4.4)	10.1 (7.3)	71.2 (48.9)	2.7** (2.9)
MSNERA	4.6 (6.0)	7.5 (8.0)	119.4* (136.0)	2.8 (2.4)

[a] Standard deviation is given in parentheses. Omissions indicate missing data or small sample size which precluded valid calculations.

[b] F female, M male, W married, S single, C children at home, N no children, E employed, U unemployed, O homeowner, R home renter, A auto, P no auto in household. Only role groups with sample sizes above twenty are shown.

* Significant at 0.05 level. ** Significant at 0.01 level.

range from 2.5 to 5.3 kilometres, and the standard deviations vary from 1.9 to 6.9 kilometres. Even though the statistics for the employed (E) subpopulation very closely match those for the total sample of respondents, its disaggregation into role-group and cluster analysis categories shows a marked divergence in all quadrants of the graph. All of the dichotomous pairs of single-variable subpopulations show high polarization in their statistics, but this is particularly strong for the automobile/nonautomobile, male/female and employed/unemployed groups. A clearer indication of the characteristics which dominate the differentiation of the space is given by the role groups. Thus, in general, the four-variable and six-variable categories indicate that males and married respondents move greater distances from home, and they tend to be more diverse in this behaviour than single respondents and females. The cluster groupings 8 (family-oriented males with high incomes) and 9 (married middle-aged blue-collar males) reinforce this interpretation. In contrast, the position of cluster grouping 2 (dominantly family-oriented, low-income housewives) shows the constraint of household obligations on mobility. Interestingly, students

Table 8. Selected average daily travel indicators for cluster analysis groupings[a].

Group	Maximum distance from home (km)	Speed for all trips ($km\ h^{-1}$)	Total time for all trips (minutes)	Distance to usual shopping location (km)
1 Students, mostly single males	2.8 (1.9)	7.9 (4.8)	104.4* (50.9)	2.4 (1.7)
2 Young, family-oriented, low-income housewives	2.7** (1.9)	4.7* (5.3)	39.0* (46.8)	2.9 (2.2)
3 Young, blue-collar, low-income singles/couples	4.4 (4.1)	9.3** (9.4)	84.2* (69.7)	2.8 (2.5)
4 Young, low-income, low mobility, low work	2.5* (4.0)	5.1* (6.4)	54.6* (46.8)	2.9 (2.7)
5 Young, highly-educated, low-income singles/couples	3.5* (3.6)	7.2* (8.3)	71.6 (43.1)	2.4* (1.9)
6 Young, high-income, small families, high mobility	4.4 (3.8)	8.9 (8.4)	75.7* (51.7)	3.5* (2.6)
7 Family-oriented housewives of average income	3.7 (5.7)	8.5 (11.5)	48.6* (50.7)	3.6* (2.7)
8 Family-oriented working males with high incomes	5.1* (4.7)	11.5* (10.0)	79.1* (43.3)	3.0 (2.9)
9 Blue-collar, middle-aged, married males	4.8* (4.8)	10.2* (10.7)	73.7 (36.8)	2.6** (1.9)
10 Mature, preretirement, mostly female	3.1* (3.8)	7.0* (9.0)	71.6* (43.1)	2.8 (2.3)

[a] Based on Ward's algorithm (Ward, 1963; Wishart, 1982). Standard deviation is given in parentheses.

* Significant at 0.05 level. ** Significant at 0.01 level.

(grouping 1) share the same restrictions on mobility, but possibly this relates to a tendency to live close to an educational institution and to the intensity of coupling constraints on student life. The greatest variability on this measure is associated with unemployed family-oriented males (MWCU). As with all of the role-group categories, it is possible to disaggregate these respondents further, to differentiate those who are retired, disabled, and temporarily out of work from those who choose not to work for other reasons; however, even a sample of more than 1200 persons does not allow for sufficient group sizes at these fine levels of disaggregation.

The average speed for all daily trips is one indicator of transportation capability. Not surprisingly, it is shown in figure 2 that the greatest polarization for single-variable subpopulations relates to the accessibility of an automobile. However, marital and employment status also show high separation on the mean and standard deviation statistics. Overall, the range in average values is from 4.7 to 12.1 km h^{-1} (kilometres per hour). At the lower end of this range are low-income groupings (2 and 4), unmarried role groups of both sexes, and those without automobiles. In general, employed and married respondents with jobs log the fastest average speeds in their daily travels and show high variability. However, the most extreme levels of within-group variation are recorded by unemployed housewives who have no children at home and by female single parents who have jobs.

Figure 3 is dominated by outliers for single, employed males with no children, who spend much more time travelling on the average day, but also show extreme variability in this measure. Single males with children also spend large amounts of time in travel, but show less variability because of the constraints of child care responsibilities. Among the single-variable subpopulations, sex, marital status, and employment status show the greatest differences. The lower left quadrant of the graph includes those groupings who devote the least of their daily time budgets to travel, and those who show the lowest within-group variation. These are dominated by unemployed respondents, mostly housewives with children to care for—roles that tend to commit one to the dwelling and to restrict outside activities.

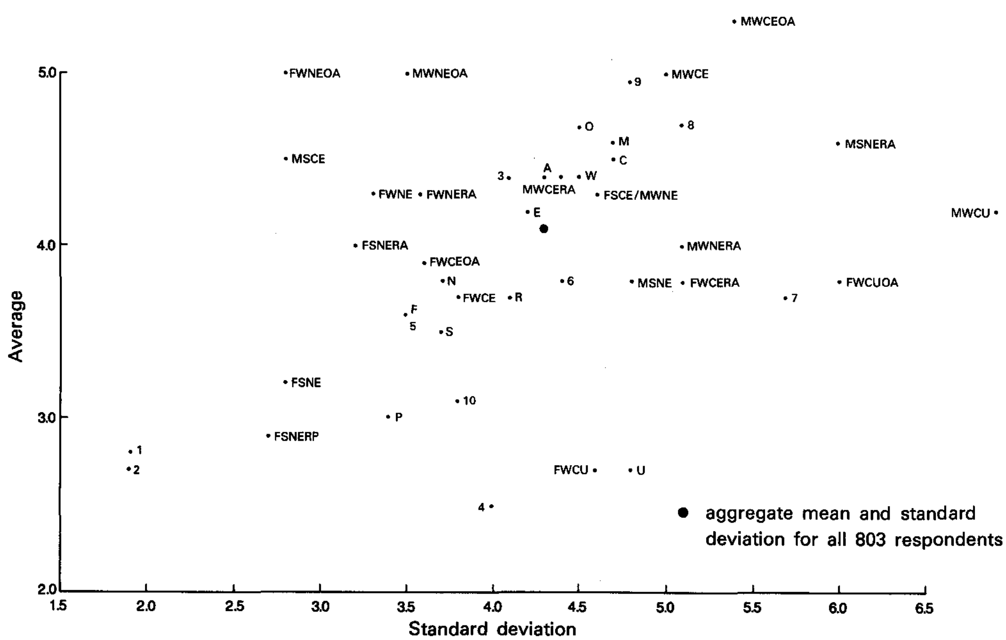

Figure 1. Maximum distance in kilometres from home on diary day for subpopulations and role groups.

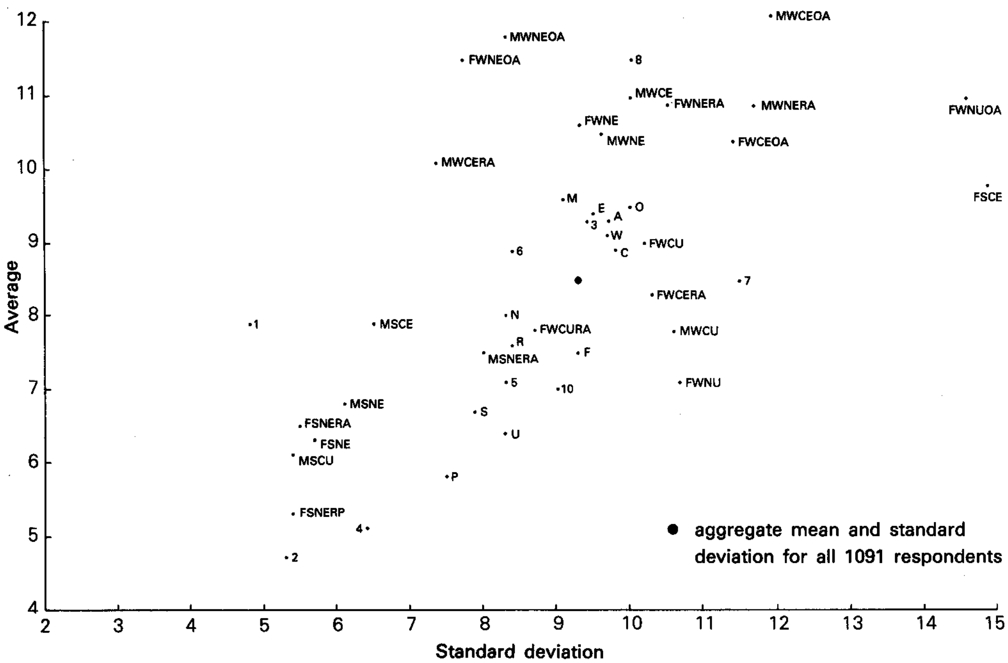

Figure 2. Average speed in kilometres per hour for all trips regardless of mode, for subpopulations and role groups.

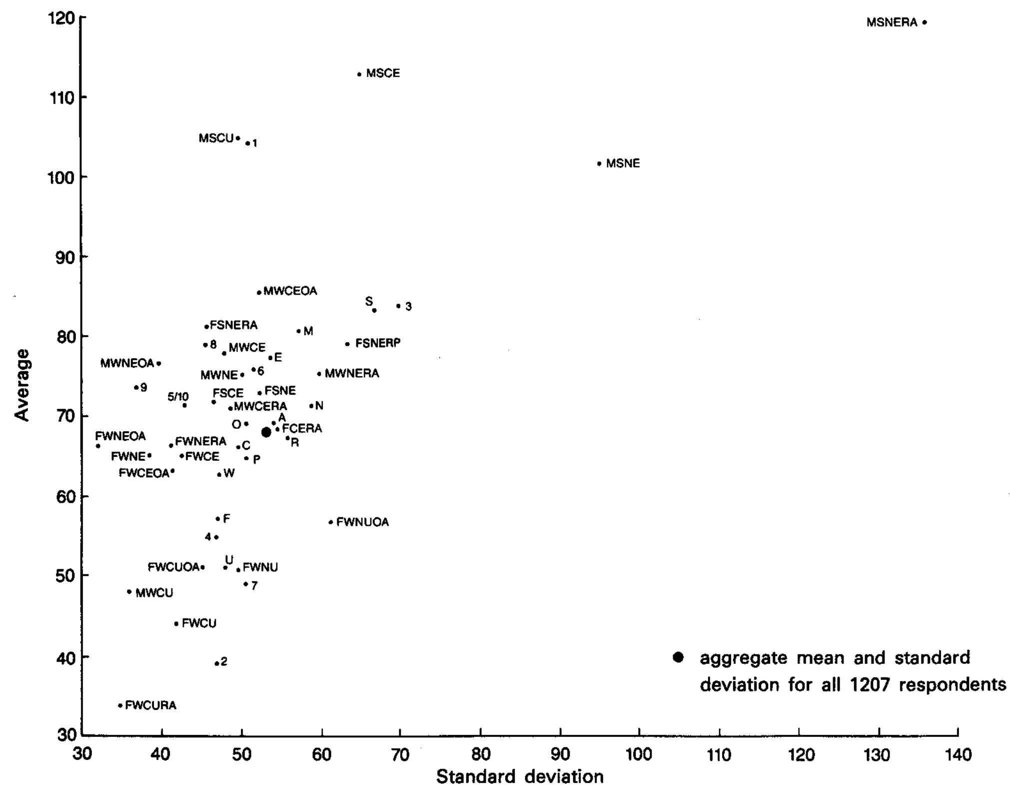

Figure 3. Average total time in minutes for all trips on diary day, for subpopulations and role groups.

The final measure is average distance to the usual shopping location. As this site is not necessarily the closest shopping opportunity for given households, it is not possible to say whether the average distance values represent preferences or geographical necessity. Because of the inevitable circularity of possible arguments over this issue, and because of difficulties in structuring a meaningful empirical framework for resolving it, we confine our interpretation to the pattern of means and standard deviations for the different methods of grouping respondents, and this is shown in figure 4.

The interpretation points to distinct cleavages based on social and demographic attributes, and to systematic relationships with the other measures of travel behaviour. Thus, groups with low values for the measures of speed and geographic range tend to use basic shopping services that are closer to their homes than do those who travel at higher speeds and over greater distances. Even though the dichotomous subpopulations based on employment have statistics similar to those of the total sample, role groups made up of unemployed respondents are dominant at the extremes of the distribution. Those with the highest means and standard deviations are married and have children to care for, whereas those who live closest to their shopping locations are single unemployed parents. All of the homeowner role groups, and all of the family-oriented groups from the cluster analysis, use shopping facilities at distances that are close to or greater than the average distance for the sample population. This is associated with the dispersed nature of suburban shopping facilities and the attraction of suburban locations for homeownership. A notable exception to this general pattern is the relatively high

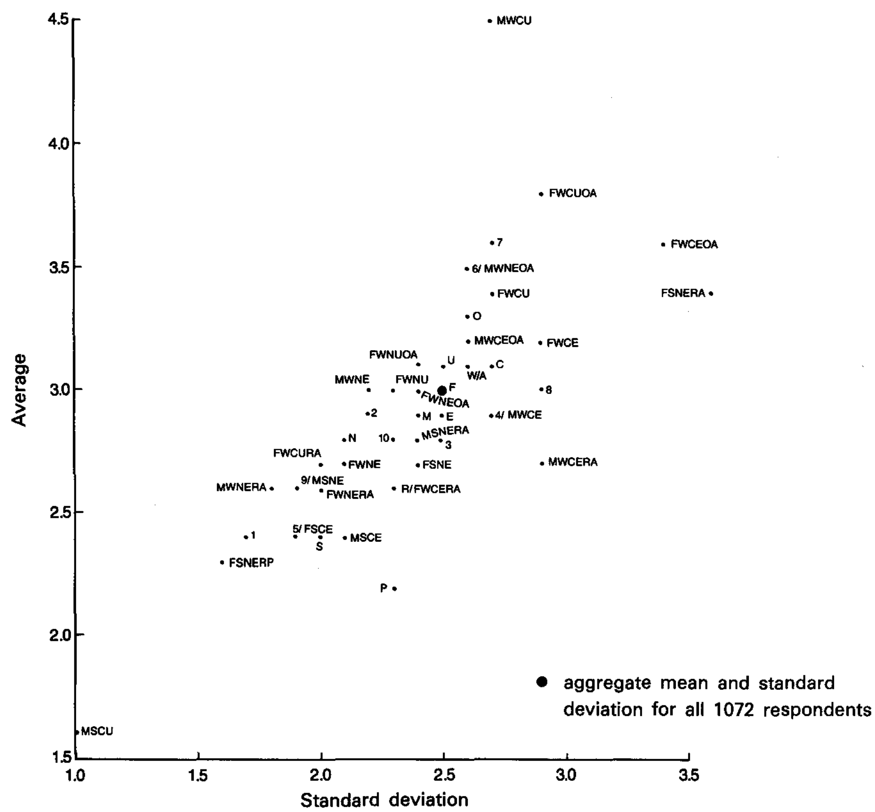

Figure 4. Average distance in kilometres to usual shopping location, for subpopulations and role groups.

mean and the very high standard deviation associated with single employed females who have automobiles (FSNERA), a group that is in sharp contrast to its non-automobile counterpart (FSNERP). These extremes emphasize the significance of mobility in the exercise of locational choice.

Conclusions

Given the similarity of personal characteristics that were included within each classification scheme, we might have expected categories which share the same traits to exhibit similar statistics on the four measures of transportation behaviour. However, the means and standard deviations varied widely for the different levels of aggregation. Nonetheless, there was little difference in the ability of each scheme to produce groupings which were homogeneous, not only in the respondents' characteristics but also for the four transportation indicators. As illustrated by the within-group's sums of squares, a priori groupings compared favourably with those based on cluster analysis. If this is the case generally then the pragmatic choice appears to favour schema based on a priori groupings.

Although cluster analysis produces groups which are 'natural' in the statistical or ecological sense, the inability to define the clusters in a definitive way is often a hindrance to their use in interpretation. In contrast, though limited to a smaller number of differentiating variables, the role-group classifications are defined with greater certainty, and this allows for easier interpretations of the effects of the combination of different variables to produce the role-group categories.

An additional reason for favouring a priori categorization methods is that they give the researcher greater control in matching the classification of respondents with the need to target specific populations, either for policy evaluations or for theoretical reasons. Whereas multivariate procedures are able to segment populations in ways which maximize, in a statistical sense, within-group homogeneity and between-group heterogeneity, there are many instances in which either a theoretically based typology or a policy-guided classification would provide a stronger foundation for analysis. Examples include the need for unambiguous groupings in research based on feminist theory and in research which addresses the ability of social policies to help targeted populations. The flexibility with which large space-time diary data sets can be disaggregated is an asset in meeting such requirements.

The methodological findings of this paper are limited to the evidence provided by a set of single-day space-time diaries. Yet, Huff and Hanson (1986) advocate caution in deriving inferences about individual behaviour patterns from single-day diaries. Their investigation of diaries kept for 35 days by respondents in Uppsala, Sweden revealed that each person followed a number of archetypical daily patterns, and that the similarity in these patterns from day-to-day was low. Although the Halifax study is based on single-day diaries, the sample was drawn to give nearly equal representation for each of the weekdays. It is clear from the results of this study that although travel behaviour may vary widely within aggregate groups, as evidenced by the high standard deviations in some instances, there are also substantial and systematic variations in behaviours between aggregate groups. However, the Halifax data does not allow us to differentiate between interpersonal variation in travel behaviours within groups or differentiate possible day-to-day variations in each individual's behaviour.

Aside from the limitation of single-day records on individual behaviour, the substantive empirical findings of this research are constrained by certain features of the study's design. A comprehensive analysis of travel behaviour would require greater representation of the contextual environment which shapes individual decisions. One would need information on the interaction effects of behaviour with the supply of transportation options, and the spatial distribution and operational timing of sites of activity. In addition, comparisons with the findings of research would require comprehensive concern for the different types of urban environment, as reflected in their size, function, social composition, and political milieu.

Although caution is in order, we do wish to draw attention to one surprising observation. It concerns the possibility that some behavioural indicators are conserved at the aggregate level across social cleavages. It is shown in table 4, for example, that whereas automobile owners travel much greater average maximum distances from home than nonautomobile owners (4.4 versus 3.0 kilometres), they spend very little more time in daily travel (68 versus 65 minutes). Similarly, although there are large differences in distance travelled to usual shopping locations between automobile owners (3.1 kilometres) and nonautomobile owners (2.2 kilometres), there are almost no differences between males (2.9 kilometres) and females (3.0 kilometres). Despite the complexity of individual travel behaviour, these observations suggest that the flexibility of space-time diary data may allow for a systematic search for behavioural constants. A parallel case has been made by Hupkes (1982).

Hupkes reported on a longitudinal analysis of travel behaviour in the Netherlands, comparing records from 1960 and 1970 on such indicators as mean distance travelled, total time devoted to travel, and trip frequency, disaggregated by mode of travel. Although the relative frequencies of various modes changed dramatically over this period, which coincided with a rapid increase in ownership of private automobiles, and although mean distances and speeds increased, both total time for travel and trip frequency tended to remain remarkably constant when they were computed as weighted averages over all modes. The combination of this kind of longitudinal analysis within the framework of space-time diaries would provide an ideal basis for the identification of those behavioural constants which are most critical in the formation of urban structures. This is an issue that has not been addressed by researchers, and this research permits, at best, only suggestive comments on the merits and utility of such an approach. Surely, in a world dominated by change, it would be useful to identify those parameters of behaviour which remain steady over significant periods of time.

This analysis has pointed to the flexibility and possible uses of large-scale space-time diary data sets. But by focusing on measures of transportation behaviour at different levels of respondent aggregation, we have drawn attention to the difficulties posed by different procedures of classification, and to the opportunities for the development of a wide array of descriptive and diagnostic measures of transportation behaviour, and for the expansion of their use in the search for behavioural constants.

Acknowledgements. We wish to thank the participants in the Halifax Time-Budget Study (Susan Clark, David Elliott, Andrew Harvey, and Dimitri Procos) for sharing their data base, the Social Science and Humanities Research Council of Canada for financial support, and the following research assistants: Bill Sawruk, Harry Taylor Jr, and Terry Van Kessel. Suggestions from Alan Hay and from an anonymous referee were appreciated.

References

Anderson J, 1971, "Space-time budgets and activity studies in urban geography and planning" *Environment and Planning* **3** 353-368

Brail K, Chapin F Jr, 1973, "Activity patterns of urban residents" *Environment and Behaviour* **5** 163-190

Burnett P, Hanson S, 1979, "Rationale for an alternative mathematical approach to movement as complex human behaviour" *Transportation Research Record* **723** 11-24

Burns L D, 1979 *Transportation, Temporal and Spatial Components of Accessibility* (Lexington Books, Lexington, MA)

Clark S, Elliott D H, Harvey A S, 1980, "Hypercodes and composite variables: simple techniques for the reduction and analysis of time budget data", paper presented at the Meeting of the International Research Group on Time Budgets and Social Activities, Sofia, 1-3 October; copy from A S Harvey, Department of Economics, St Mary's University, Halifax, Canada B3H 3C3

Cullen I, Godson V, 1975, "Urban networks: the structure of activity patterns" *Progress in Planning* **4** 1-96

Dangschat J, Droth W, Friedrichs J, Kiehl K, 1982, "Action spaces of urban residents: an empirical study in the region of Hamburg" *Environment and Planning A* **14** 1155-1174

Elliott H, Harvey A, Procos D, 1976, "An overview of the Halifax time-budget study" *Society and Leisure* **3** 145-159

Hägerstrand T, 1974, "The impact of transport on the quality of life", in *Transport in the 1980-1990 Decade. Proceedings of the 5th Symposium of the European Conference of Ministers of Transport, Athens, 22-25 October 1973* copy from T Hägerstrand, Department of Geography, University of Lund, Lund, Sweden

Hanson S, 1982, "The determinants of daily travel-activity patterns: relative location and sociodemographic factors" *Urban Geography* **3** 179-202

Hanson S, Hanson P, 1980, "Gender and urban activity patterns in Uppsala, Sweden" *Geographical Review* **70** 291-299

Hanson S, Hanson P, 1981, "The travel activity patterns of urban residents: dimensions and relationships to sociodemographic characteristics" *Economic Geography* **57** 332-347

Hanson S, Huff J, 1986, "Classification issues in the analysis of complex travel behaviour" *Transportation* **13** 271-293

Harvey A S, Elliott D H, Procos D, 1977, "Sub-populations relevant to the study of the use of time", Working Paper, Institute of Public Affairs, Dalhousie University, Halifax, Nova Scotia B3H 4H2, Canada

Hensher D A, McLeod P B, 1977, "Towards an integrated approach to the identification and evaluation of the transport determinants of travel choices" *Transportation Research* **11** 77-93

Horton F E, Hultquist J F, 1971, "Urban household travel patterns: definition and relationship to household characteristics" *East Lakes Geographer* **7** 37-48

Huff O, Hanson S, 1986, "Repetition and variability in urban travel" *Geographical Analysis* **18** 97-114

Hupkes G, 1982, "The law of constant travel time and trip-rates" *Futures* **14** 38-46

Janelle D G, Goodchild M F, 1983, "Transportation indicators of space-time autonomy" *Urban Geography* **4** 317-337

Kutter E, 1973, "A model for individual travel behaviour" *Urban Studies* **10** 235-258

Lenntorp B, 1978, "A time-geographic simulation model of individual activity programmes", in *Human Activity and Time Geography. Volume 2* Eds T Carlstein, D N Parkes, N J Thrift (Edward Arnold, London) pp 162-180

Lindsay I J, 1979 *An Exploratory Time-Space Model of Activity Selection and Travel Behaviour with Special Reference to New Forms of Work Scheduling* PhD thesis, University of Toronto, Toronto, Ontario M5S 1A1, Canada

Oppenheim N, 1975, "A typological approach to individual urban travel behaviour prediction" *Environment and Planning A* **7** 141-152

Palm R, Pred A, 1974, "A time-geographic perspective on problems of inequality for women", WP-26, Institute of Urban and Regional Development, University of California, Berkeley, CA 94720, USA

Pas E I, 1983, "A flexible and integrated methodology for analytical classification of daily travel-activity behaviour" *Transportation Science* **17** 405-429

Pas E I, 1984, "The effect of selected sociodemographic characteristics on daily travel-activity behaviour" *Environment and Planning A* **16** 571-581

Szalai A (Ed.), 1972 *The Use of Time: Proceedings of the World Time-Budget Study* (Mouton, The Hague)

Ward J, 1963, "Hierarchical grouping to optimize an objective function" *Journal of the American Statistical Association* **58** 236-244

Wishart D, 1982 *Supplement, CLUSTAN User Manual* third edition, Program Library Unit, University of Edinburgh, Edinburgh EH8 9YL, Scotland