Department for Transport

National Travel Survey 2011 GPS pilot: summary analysis

January 2012

The Department for Transport has actively considered the needs of blind and partially sighted people in accessing this document. The text will be made available in full on the Department's website in accordance with the W3C's Web Content Accessibility Guidelines. The text may be freely downloaded and translated by individuals or organisations for conversion into other accessible formats. If you have other needs in this regard please contact the Department.

Department for Transport
Great Minster House
33 Horseferry Road
London SW1P 4DR
Telephone 0300 330 3000
Website www.dft.gov.uk

Contents

Executive summary

1. This report summarises the results of the 2011 National Travel Survey (NTS) GPS pilot survey and compares them to results from the travel diary data collected by the main NTS for the same fieldwork period. There are two supplementary technical reports to this document which detail the fieldwork and data processing for this project.

2. There were no major problems with fieldwork that could not have been overcome had we more time or devices available in which to complete the project. Practical issues included an update for the MGE device management software not being available in time for use and two variables (DIST and speed) were not recorded correctly because the devices settings were not correct, however, these were later estimated using position data during the data processing stage.

3. There are a number of disparities between the GPS personal travel data and the data from the NTS diary:

   On average, seven per cent of GPS pilot respondents made journeys but forgot to carry the device with them for some or all of the trips made on each day of the travel week and three per cent of respondents reported having forgotten to charge their GPS device each day of the travel week;

   There were far fewer trips and stages in the GPS data, but the time taken to complete these was much longer and the average journey length was also much longer;

   Although the data processing confirmed that the use of combined GPS and accelerometer data improves the mode prediction, there were fewer walking trips in the processed GPS data and more rail trips;

   The distribution of GPS trips tailed off towards the end of the travel week, and

   The GPS trip distribution does not contain the traditional morning and afternoon rush hour peaks, but instead peaks once in the early afternoon.

4. The GPS data contains a greater share of trips to or from home than the diary. It also estimated fewer journeys to/from work and there were a number of substantial differences in the proportion of trips coded purpose to/from shopping trips, visiting friends or family and holidays. One quarter (25 per cent) of trips were missing either a 'to' or 'from' purpose code.

5. Questions asked in the pick-up interview on whether the respondent was the driver or a passenger in their last car trip and car occupancy are unlikely to provide a comparable alternative to the stage data collected by the travel diary, however, it may be feasible to use interview questions as a substitute for diary data on taxi usage.

6. The differences between results for the GPS sourced trip data and the diary trip data lead us to conclude that when used in the context of the NTS methodology and the Trace Annotator processing system, GPS devices do not provide an acceptable alternative to the paper travel diary. It is apparent that trips derived using GPS processing are not quite the same as the trips recorded in the diary. GPS derived trips are possibly more akin to a series of diary trips.

1. Introduction

1.1 In November 2010, the Department for Transport (DfT) contracted the National Centre for Research (NatCen), the current contractor for the National Travel Survey (NTS) to undertake a pilot survey using accelerometer equipped GPS devices to collect personal travel data in place of the paper travel diary typically used for this purpose. This project was the culmination of two earlier studies which concluded that personal GPS devices were the most suitable option for delivering affordable and practical improvements to NTS data collection, have real promise for use within the GB NTS and do not pose any fundamental barriers of feasibility or public acceptability[1,2].

1.2 A subsample of the February and March 2011 NTS sample was randomly allocated to the GPS pilot. This sample was designed with the aim of collecting GPS data from 1,000 respondents aged 12 or older over the two waves of fieldwork and consisted of 902 households. In total interviews were achieved with 1,269 individuals. Of these, 1,074 were aged 12 years or older and were asked to carry a GPS device for a seven day 'travel week': 912 agreed to carry a device, and 897 subsequently carried and returned a device. Data collection took place between 11 February and 8 June 2011.

1.3 The fully productive household response rate for the GPS pilot (52 per cent) was lower than for the main NTS during the first quarter of 2011 (59 per cent). To be classed as a 'fully productive' household, all eligible household members must complete a travel diary (or carry and return a GPS device). The main reason for the difference in response rates was that within responding households, individuals were slightly less likely to agree to carry GPS devices than complete paper travel diaries: 84 per cent of GPS pilot respondents agreed to carry and returned a GPS device, whereas 91 per cent of Q1 2011 NTS respondents completed and returned a travel diary. Interviewers noted that there were some

[1] *National Travel Survey GPS Feasibility Study Final Report*, December 2009 by Tracy Anderson, Varunie Abeywardana, Jean Wolf and Michelle Lee
http://webarchive.nationalarchives.gov.uk/+/http://www.dft.gov.uk/pgr/statistics/datatablespublications/personal/methodology/ntsreports/ntsgpsstudy.pdf
[2] *Review of the Potential Role of 'New Technologies' in the National Travel Survey* by Jean Wolf, Peter Bonsall, Marcelo Oliveira, Lauren Leary and Michelle Lee
http://webarchive.nationalarchives.gov.uk/+/http:www.dft.gov.uk/pgr/statistics/datatablespublications/personal/methodology/ntsreports/ntsreviewtechnologies.pdf

concerns from respondents regarding the devices and that the hardest respondents to persuade to participate were those who did not like or feel comfortable using modern technology, or were suspicious about the devices themselves and the idea that they were being traced.

1.4 A team at the Eindhoven University of Technology (TU/e) led by Professor Harry Timmermans, comprising of Joran Jessurun, Anastasia Moiseeva and Tao Feng were sub-contracted to process the GPS data, building on their GPS 'Trace Annotator,' using accelerometer data to improve the prediction of mode of travel[3].

1.5 There are two supplementary reports that provide a detailed account of the methodology, fieldwork and data processing of the NTS GPS pilot:

- *National Travel Survey 2011 GPS Pilot*, a technical report on the pilot survey management and data collection, by Josi Rofique, Alun Humphrey and Caroline Killpack of NatCen, and
- *Processing of National Travel Survey GPS Pilot Data*, a technical report prepared on behalf of the Department for Transport by Tao Feng, Anastasia Moiseeva and Professor Harry Timmermans at TU/e.

1.6 These will be published alongside this report on the NTS home page under the section entitled 'future developments' at http://www.dft.gov.uk/statistics/series/national-travel-survey

1.7 The terms 'trip' and 'journey' are used interchangeably in this report. A trip is defined as 'a one-way course of travel with a single main purpose'. A trip consists of one or more stages. A new stage is defined when there is a change in the form of transport or when there is a change of vehicle requiring a separate ticket.

1.8 The GPS pilot project was conducted to strict deadlines to fit into the NTS re-procurement timetable which required work be completed by December 2011. The project start was delayed for several months following the change of Government in May 2010, placing considerable pressure on this deadline. More work could have been done to investigate and potentially re-process some of the unusual results in the GPS data (for example, the presence of walking stages of 50-99 miles in length) had more time been available.

[3] *Semi-automatic Imputation of Activity Travel Diaries: Use of Global Positioning System Traces, Prompted Recall, and Context-Sensitive Learning Algorithms*, Anastasia Moiseeva, Joran Jessurun, Harry Timmermans, Eindhoven University of Technology. Submitted to the Transport Research Board 2010 http://trid.trb.org/view.aspx?id=909867

2. Headline results from processed GPS data

2.1 Please note that the stage and trip analyses presented in this report use the data set as delivered to the department: cases have been included regardless of whether the length and mode, purpose or start/end time were feasible. After the data processing was completed, six expert-set, rule based algorithms were applied to the results to check whether trips appeared plausible, specifically:

- The distance and duration of trips with respect to the mode used;
- The distance travelled with respect to the trip purpose, and
- The start time, end time and duration of a trip in relation to the trip purpose.

2.2 This work concluded that approximately 5-10 per cent of cases were 'implausible' or 'possible, but less plausible'[4]. These cases were not re-processed, flagged or removed from the data and therefore feature in the analysis within this report.

Missing data

2.3 There are a number of reasons why a respondent carrying a GPS device may not have recorded data: they may not have made any journeys; they may have forgotten to charge the device and it may subsequently have run out of power, or they may have made journeys but forgot to carry the device with them. Respondents were asked during the pick-up interview whether there were any days when such incidents occurred.

2.4 On average, on each day of the travel week approximately 14 per cent of people said they did not make any journeys, three per cent said they forgot to charge the device and seven per cent reported having forgotten to carry the device. Around half the respondents said they had travelled every day (46 per cent). The majority reported having remembered to

[4] These being trips and other 'activity episodes' that featured in the Trace Annotator processing but do not relate to travel activity.

charge their device every day (85 per cent) and around three quarters said they carried the device with them every day (73 per cent).

Table 2.1: Number and proportion of respondents who did not travel, charge device, or made journeys and forgot to carry the device (by day)

Day of travel week	No journey	No charge	No carry
Day 1	13%	2%	7%
Day 2	13%	3%	7%
Day 3	15%	3%	7%
Day 4	14%	3%	6%
Day 5	15%	3%	6%
Day 6	14%	3%	6%
Day 7	12%	3%	8%
Travelled every day	46%	-	-
Didn't forget to charge the device	-	85%	-
Didn't forget to carry the device	-	-	73%
Can't remember	13%	8%	7%
Total	100%	100%	100%

Note: all respondents (871: excludes 26 missing/ 'don't know' responses).

2.5 The GPS devices used in the pilot record information on battery power levels and it is possible to identify whether a device has run out of power and thereby stopped functioning. However, TU/e have not provided any information on the frequency of such events.

2.6 By summing the number of people who said they made trips but forgot to carry the device on days 1-7, we can conclude that there were in total, 403 travel days were missing from the data set. There were potentially a total of 6,279 travel days in our data set, so we can conclude that data was missing for approximately seven per cent of travel days. However, devices were not necessarily forgotten for all trips in a day: three per cent of stages and three per cent of journeys were made on days when respondents reported having forgotten to carry the device. It is also possible that due to poor recall, respondents may have incorrectly

reported having forgotten the device on a particular day when they did carry it.

Stages

2.7 There were 12,515 stages in the processed GPS data. Of these, 536 were designated as 'waiting' activity at bus stops or rail stations between stages involving travel activity. The following analysis only concerns the 11,979 stages concerning travel activities. These stages accounted for 276,512 minutes of travel time and a total distance of 567,065 miles. The mean length of stages was 23 miles or 47 minutes.

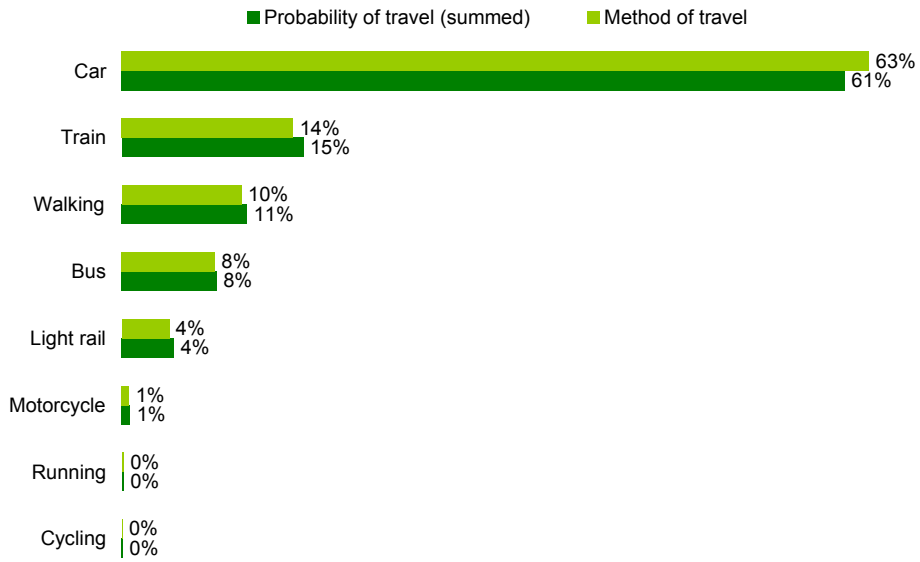
Distribution of stages

Mode of transport

2.8 The majority of GPS stages were made by car (63 per cent), 14 per cent were made by train, 10 per cent by walking, eight per cent by bus and four per cent by light rail (tram, DLR, tube etc.). Just one per cent of stages were made by motorcycle, and less than one per cent of stages involved cycling or running.

2.9 The Trip Annotator used to infer the mode, accorded each mode of transport a probability value. On this basis the mode of transport can be analysed using two different methods: by allocating each stage a single mode of transport (that allocated the highest probability) or by summing the probabilities awarded to each mode. Figure 2.2 contains the results for both of these methods. As one would expect, the two produce similar results, however in terms of distribution, the summation of probabilities tends to allocate slightly fewer cases to the most dominant form of transport (car).

Figure 2.2: Distribution of stages by mode



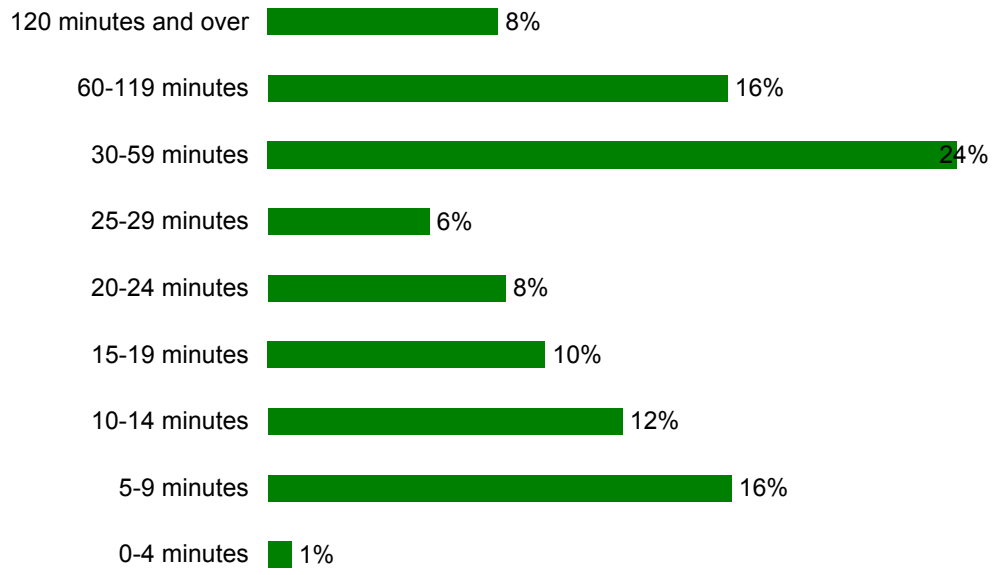
Base: 11,979 stages.

2.10 Whilst the summation of probabilities might be regarded as a superior way of analysing mode of transport - it accounts for all the cases where non-dominant modes are assessed as being nearly as likely as the dominant mode and thereby reduces the overall error of prediction - it is not readily practicable for use with a standard data analysis tool which requires each case within the data set to be allocated a fixed value for each variable. The summation technique works when summing across all cases or and in cross-tabulations, but it cannot be used to filter cases or derive new variables according to mode.

Length of stage

2.11 As noted in paragraph 2.7: the mean time taken to complete a stage was 47 minutes, however the modal value was five minutes and the median value was 27 minutes. This suggests that a small number of high values may be skewing the mean result (indeed the maximum trip length was 846 minutes). Approximately half of all stages (52 per cent) took less than 30 minutes to complete and around a quarter took at least 30 minutes but less than an hour (24 per cent). Sixteen per cent of stages took more than one hour but less than two hours to complete and eight per cent took two hours or longer to complete.

Figure 2.3: Length of stages (minutes)



Base: 11,979 stages.

2.12 Table 2.4 outlines how stage length varies according to mode. Walking stages tended to be shorter than stages using other modes: 59 per cent took less than 20 minutes to complete and only one in ten (10 per cent) took longer than an hour. In contrast, rail took longer than other modes: only 28 per cent of train stages took less than 20 minutes to compete and nearly a third (32 per cent) took at least an hour to complete.

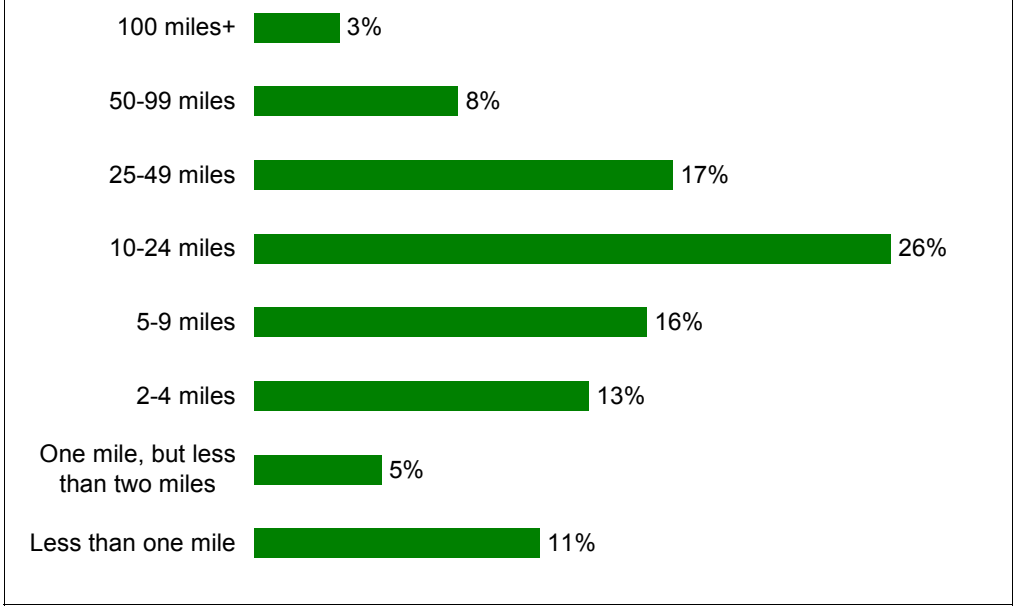
Table 2.4: Trip length for selected modes (minutes)

	Walking	Bus	Car	Train	Light rail	All modes
0-9 minutes	23%	17%	17%	13%	11%	17%
10-19 minutes	36%	20%	21%	15%	24%	22%
20-29 minutes	18%	12%	13%	14%	14%	14%
30-59 minutes	12%	25%	25%	26%	26%	24%
60-119 minutes	7%	18%	15%	22%	17%	16%
120 minutes and over	3%	8%	8%	10%	8%	8%
Median duration	16	31	28	37	30	27

Note: excludes running, cycling and motorcycle owing to small number of stages allocated to this mode.
Base: 11,878 stages

2.13 While the mean stage length was 23 miles, the modal value was three miles and the median value 12 miles: again the mean may be skewed by extreme outliers (the maximum trip length was 588 miles). Eleven per cent of stages were less than one mile long, 16 per cent were less than two miles long and 30 per cent were less than five miles long. Nearly one quarter of stages were 10-24 miles long (26 per cent), 17 per cent were 25-49 miles long and 11 per cent were 50 miles or longer.

Figure 2.5: Length of stages (miles)



Base: 11,979 stages.

2.14 The majority of walking trips were less than two miles in length (81 per cent) and 94 per cent were less than 10 miles long. Only 5-10 per cent of stages using bus, car, train or light rail were less than two miles long. Approximately one in three stages completed using bus, car or light rail were 2-9 miles in length (34 per cent, 32 per cent and 34 per cent respectively) and a similar proportion were 10-24 miles long (30 per cent, 27 per cent and 30 per cent). Rail trips tended to be longer than stages using other modes, with 41 per cent of these being at least 25 miles long.

Table 2.6: Trip length for selected modes (miles)

	Walking	Bus	Car	Train	Light rail	All modes
Less than two miles	81%	10%	10%	5%	8%	17%
2-9 miles	13%	34%	32%	22%	34%	29%
10-24 miles	3%	30%	27%	31%	30%	26%
25-49 miles	2%	17%	18%	24%	17%	17%
50-99 miles	1%	6%	9%	13%	9%	8%
100 miles+	.	3%	4%	5%	3%	3%
Median distance	1	12	13	21	13	12

Note: excludes running, cycling and motorcycle owing to small number of stages allocated to this mode.
Base: 11,878 stages

Day of the travel week

2.15 NTS travel week start days are allocated so that the choice of the travel week in which the travel diary is completed, is not left to the discretion of the respondent or the interviewer as this could introduce bias. The NTS is a continuous survey for which the annual reference year is split into twelve quota months which typically run from the middle of the month. Interviewers are given a travel week allocation card which lists a randomly selected travel week start date for each interview they need to undertake.

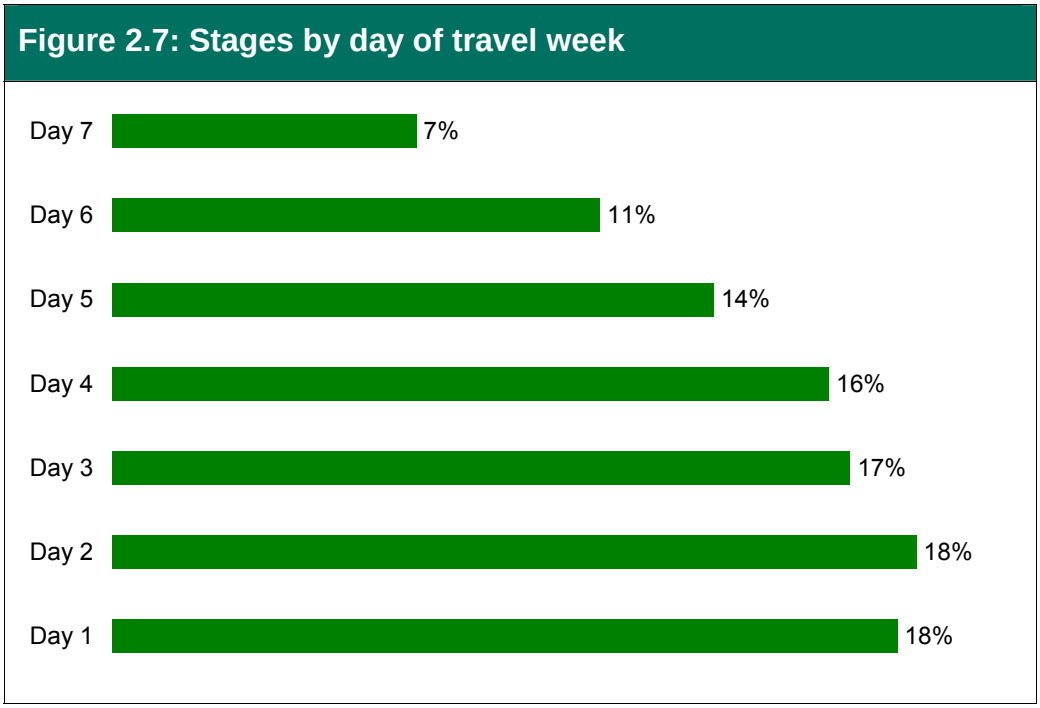
2.16 The GPS devices used in the pilot study were leased at a daily rate. Thus it was not efficient to allocate travel days in the usual method, whereby they might not be in use for several days or even weeks. Instead interviewers were advised to undertake interviews (and thereby start travel weeks) as soon as sufficient devices were available. This

decision was made even more necessary when the fieldwork start date was delayed because the devices were not delivered on time. It was originally intended that the GPS pilot fieldwork would start in January 2011. However, it was delayed until mid-February when approximately half of the required devices became available, with subsequent deliveries later in the month. The limited number of devices meant that interviewers had to allocate travel weeks as and when sufficient devices were available. This may mean that some bias was introduced to the pilot in the travel week start day.

2.17 In addition to this, the daily availability of GPS devices may not have been uniformly distributed because of practical considerations. For example, devices would not be processed or redistributed at weekends. It is also possible that likely that given the choice, NTS interviewers prefer not to interview on Sundays (meaning that GPS travel weeks could be slightly less likely to start on a Monday).

2.18 Figure 2.7 suggests that the number of stages recorded on each day of the travel week declined towards the end of the travel week. Each of the first four days of the travel week account for 16 -18 per cent of stages. This falls to 14 per cent on day five, 11 per cent on day six and seven per cent on day seven.

Figure 2.7: Stages by day of travel week



Base: all cases (11,979)

2.19 An obvious reason for this might be that respondents forget to charge or carry their devices as their motivation waned although Table 2.1 suggests that this was not the case. If we were to assume that the drop off through the week is caused by people forgetting to charge or carry the device (for the whole day) we could apply the average rate for days 1-4 to days 5-7 and estimate that there would be a total 14,380 stages: equivalent to 836 stages per person per year, had respondents carried the device with them at all times.

Other categories of stages

Short walks

2.20 In order to reduce the burden on respondents the instructions for completing the NTS paper travel diary only require reporting of short walks (those of at least 50 yards, but less than one mile) on day seven of the travel week. These are then multiplied by a factor of seven. This weighting does not apply to GPS data, where (in theory) all trips including short walks are recorded. Applying the short walk definition to the GPS data results in six per cent of GPS stages being categorised as walks of less than one mile in length. This selection is based upon the allocation of one single mode of transport to each stage. In contrast, 21 per cent of (weighted) diary stages were short walks.

Off-network stages

2.21 Travel activities off the transport network do not fall within the remit of the DfT, therefore the NTS travel diary instructs respondents to report any trips of these nature (for example, off -road mountain biking or walking in the countryside). As noted in the previous paragraph, any ineligible trips of this nature would automatically be recorded by a GPS device. TU/e were asked to create a filter that would identify stages where travel activities were undertaken off the transport network. They therefore flagged stages travelled by car (based upon the allocation of one single mode of transport to each stage) as 'off network' when the average value of the distance between the nearest road and the GPS trace was at least 10 meters. On this basis, four per cent of car stages were defined as 'off network'.

2.22 In practice the transport network or 'public highway' is not just roads but also includes public footpaths and cycle paths. Thus off-network travel could also include active travel including walking, running and cycling, and motorcycles. Realistically, off-network travel activities are likely to be part of a stage which combines on and off network travel and in turn further hinders the allocation of such a definition. Consider for example, a

person running to work: the majority of the trip is likely to be on public footpaths but sections of it may run through parks or private roads.

2.23 Although these stages are not recorded by the NTS, they were not filtered from the GPS data. Firstly, only 317 stages were classified as 'off network' suggesting that the overall impact of this these stages on top line values would be relatively minor. Furthermore, we had originally anticipated that such a filter would include a broader range of activities and were also concerned how this filter could be applied to both stages and trips. Journeys would be identified as 'off network' if the main stage (i.e. longest distance stage) were travelled by car and described as off-road. That is, a journey defined as off network would not actually be entirely off road.

Journeys[5]

2.24 In total the processed GPS pilot data contained 11,443 journeys, 611 of which would be considered short walks under the NTS definition. These accounted for 276,512 miles or 581,900 minutes of travel. The GPS data processing produced similar volumes of trips and stages (a total of 11,443 trips and 12,515 stages). From this we can conclude that the majority of trips were single stage trips, which in turn means that the distribution of journeys is similar to that of stages.

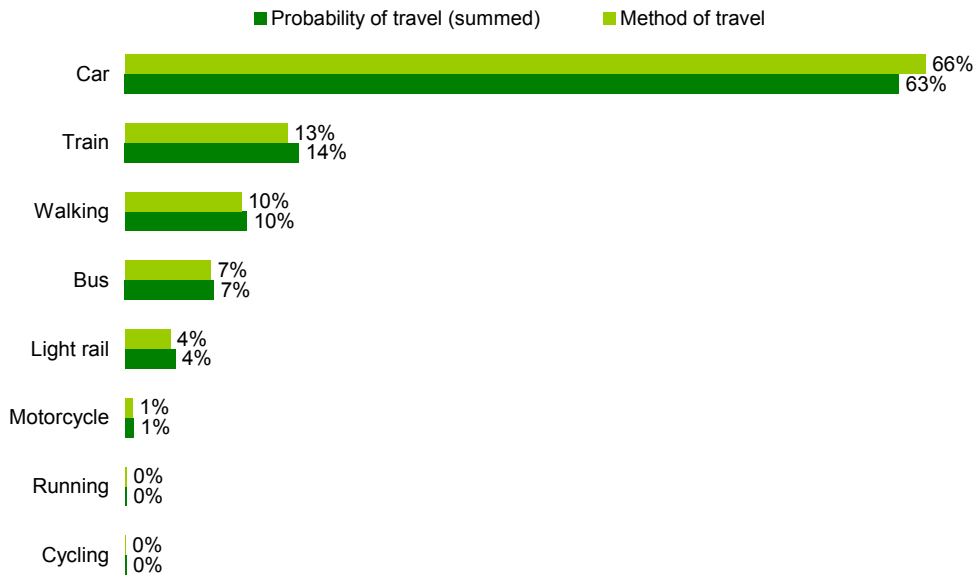
2.25 The following analysis is based on 11,090 trips, 606 of which were short walks. These exclude trips relating to children playing in the street and trips undertaken by those transporting people or goods during work hours. They account for 265,073 miles and 553,768 minutes. The mean journey was 24 miles long and took 51 minutes to complete.

[5] We do not discuss the relationship between stages and journeys here as it is difficult to filter the two on a comparable basis. Whilst waiting stages are excluded from the stage data, the trip data also excludes children playing in the street - a journey level variable that would need retro-fitting to the stage data - and work done transporting people or goods - a filter based upon trip start and end times, which are not present in the stage data.

Main mode of journey

2.26 The main mode[6] used for the majority of trips of journeys was car (66 per cent). Approximately one in ten trips used rail or walking for the main mode (13 per cent and 10 per cent respectively) seven per cent used bus, four per cent light rail and one per cent used motorcycle for the main mode.

Figure 2.8: journeys by mode



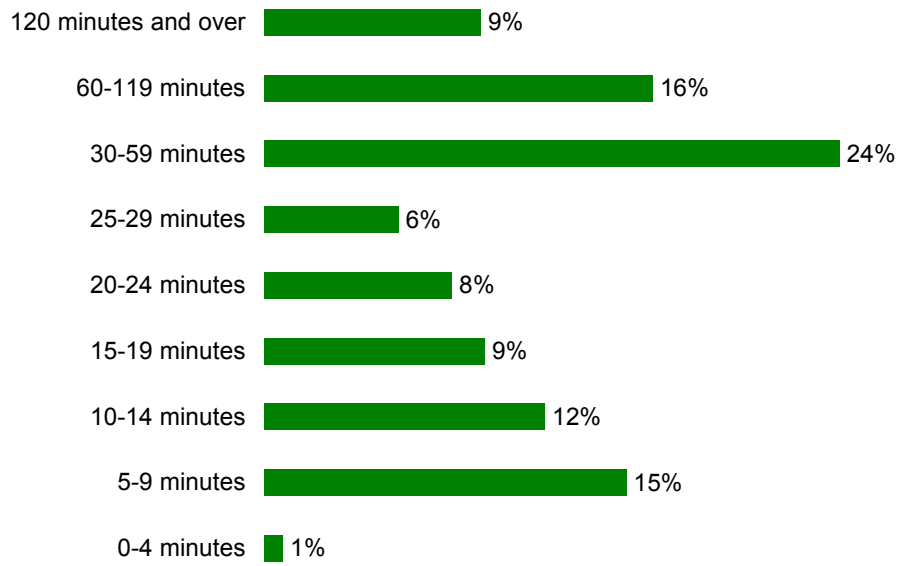
Base: all journeys excluding children playing in the street and trips transporting people or goods during working hours (11,090).

[6] Main mode of a journey is determined as the mode which was used for the longest distance stage.

Length of journey

2.27 The median journey length was 29 minutes and the mode value was five minutes. Figure 2.9 illustrates the distribution of time taken to complete a journey.

Figure 2.9: journeys length (minutes)



Base: all journeys excluding children playing in the street and trips transporting people or goods during working hours (11,090).

2.28 As with stages, trips where the main mode was walking tended to be shorter than trips involving vehicles or rail. Seventy-one per cent of the trips where the main mode was walking took less than 30 minutes to complete and 85 per cent took less than an hour to complete (Table 2.10).

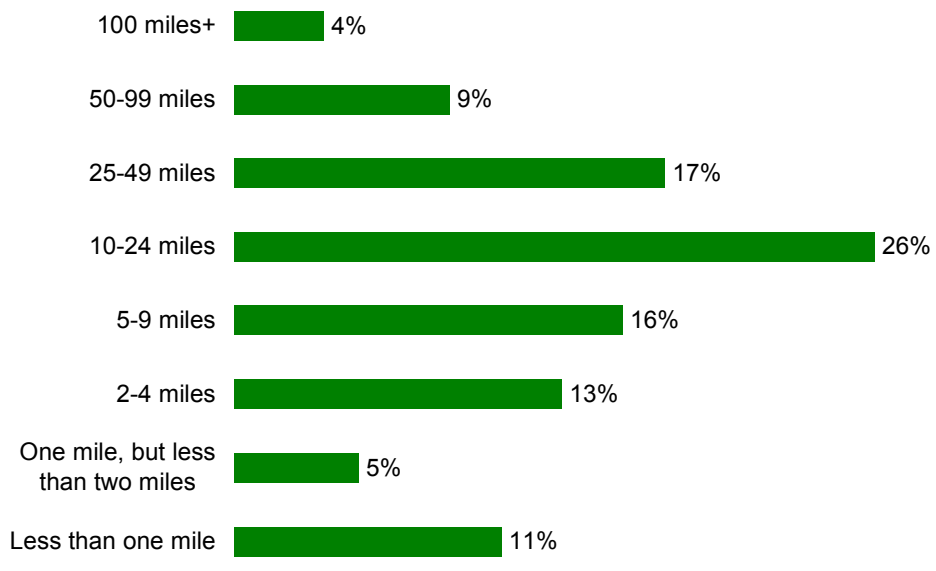
Table 2.10: Journey length (minutes) by selected modes

	Walking	Bus	Car	Train	Light rail	All modes
0-9 minutes	21%	14%	17%	11%	10%	16%
10-19 minutes	33%	17%	21%	13%	24%	21%
20-29 minutes	17%	10%	13%	13%	14%	13%
30-59 minutes	14%	26%	25%	27%	22%	24%
60-119 minutes	9%	19%	15%	23%	19%	16%
120 minutes and over	6%	13%	8%	13%	11%	9%
Median duration	17	38	28	42	32	29

Note: excludes cycling, running and motorcycle due to small number of cases
Base: 10,997 trips.

2.29 The median journey distance was 12 miles and the modal value was three miles. Figure 2.11 shows how journey distance was distributed.

Figure 2.11 Journey length (miles)



Base: all journeys excluding children playing in the street and trips transporting people or goods during working hours (11,090).

2.30 Again, trips where walking was the main mode tended to be shorter than for other modes and trips where rail was the main mode tended to be slightly longer than with other modes. Nearly four in five trips where walking was the main mode were less than two miles in length (78 per cent). In contrast, 39 per cent of trips where rail was the main mode were at least 25 miles long.

Table 2.12: Journey length (miles) by selected modes

	Walking	Bus	Car	Train	Light rail	All modes
Less than two miles	78%	9%	10%	4%	8%	16%
2-9 miles	15%	30%	32%	19%	32%	29%
10-24 miles	4%	30%	27%	31%	27%	26%
25-49 miles	2%	19%	18%	24%	19%	17%
50-99 miles	1%	8%	8%	15%	10%	9%
100 miles+	.	4%	4%	6%	4%	4%
Median distance	1	15	13	22	14	12

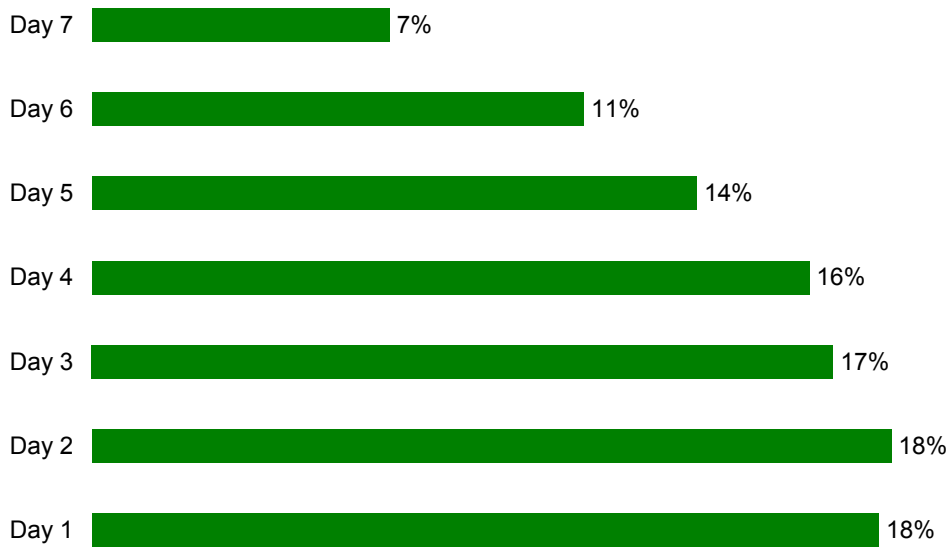
Note: excludes cycling, running and motorcycle due to small number of cases
Base: 10,997 trips.

2.31 The differences between the average stage and trip lengths and the median stage and trip lengths in both miles and minutes indicate that there are some substantial outliers present within the data. It is possible that given more time to undertake further checking and validation, some of these values could be excluded, possibly producing closer results.

Day of travel week

2.32 As with stages, the number of trips recorded each day declines as the travel week progresses (Figure 2.13). Days 1-4 accounted for approximately 16-18 per cent of trips, however, the daily share of trips fell to 14 per cent on day five, 11 per cent on day six and seven per cent on day seven.

Figure 2.13: journeys by day of travel week



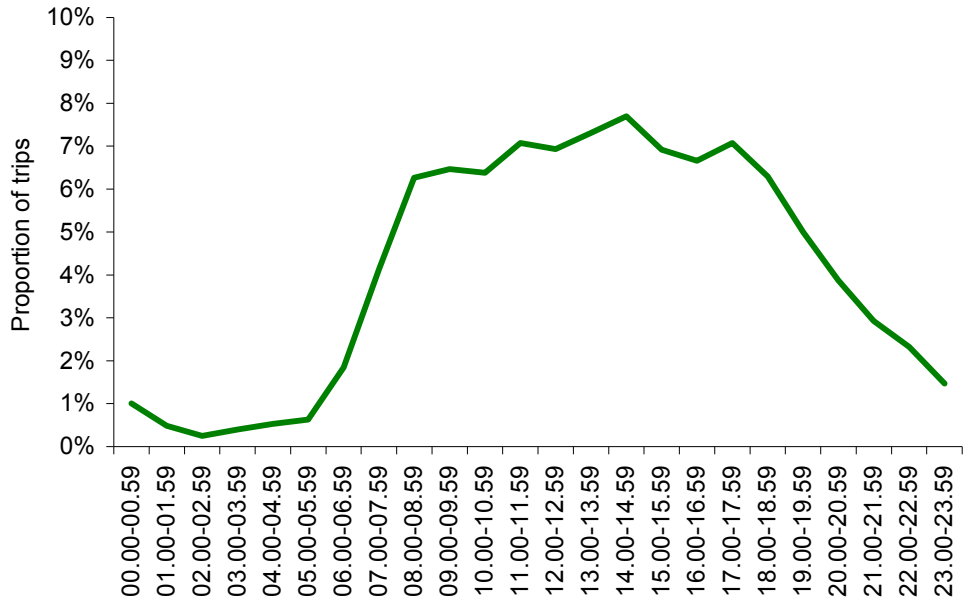
Base: all journeys excluding children playing in the street and trips transporting people or goods during working hours (11,090).

2.33 If we were to assume that the drop off through the week is caused by people forgetting to charge or carry the device (for the whole day) we could apply the average rate for days 1-4 to days 5-7 and estimate that there would be a total 13,312 trips: equivalent to 774 trips per person per year, had respondents carried the device with them at all times.

Start time of journey

2.34 Figure 2.14 illustrates the distribution of journeys by start time. Unsurprisingly perhaps, three quarters of all journeys take place during the working day between 08.00-18.59. The number of journeys started peaks just after lunch between 14.00-14.59 (eight per cent).

Figure 2.14: journeys by start time

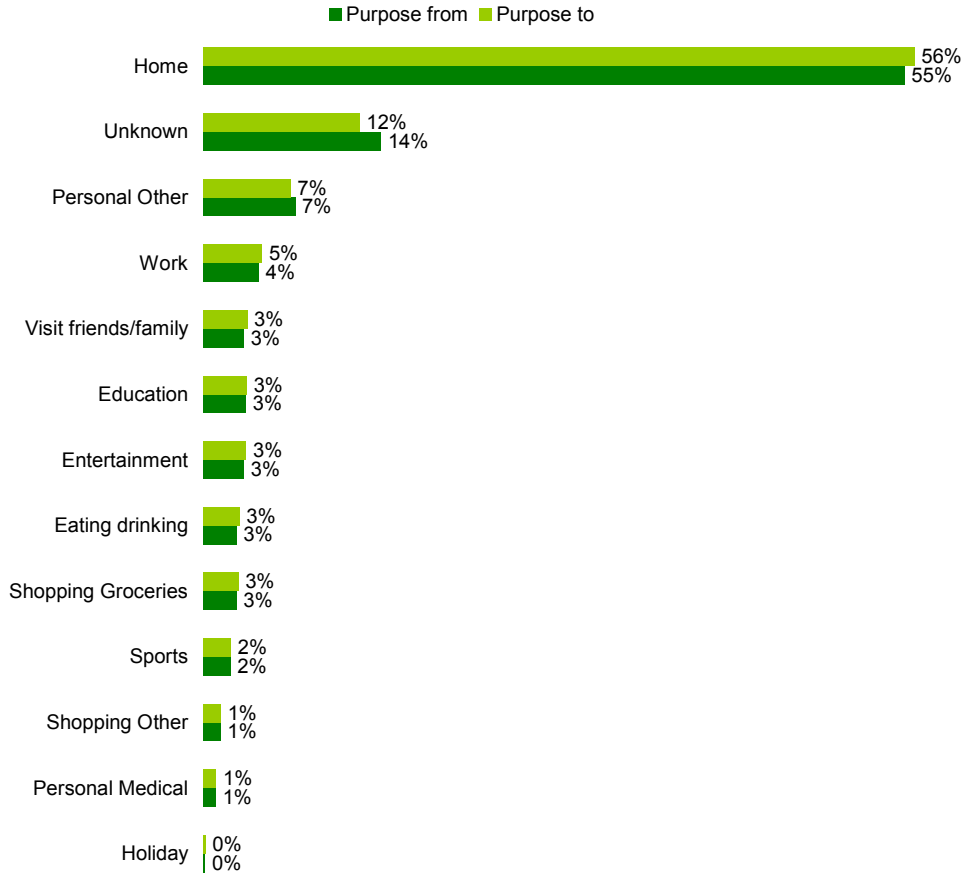


Base: all journeys excluding children playing in the street and trips transporting people or goods during working hours (11,090).

Purpose of journey

2.35 The most common purpose for journeys was travelling to or from home (56 per cent and 55 per cent respectively). Twelve per cent of journeys to a destination were coded 'unknown' as were 14 per cent of journeys from an origin. The remaining purpose categories accounted for relatively small shares of trips. Seven per cent of trips were coded to/from other personal business. This category was coded using GIS Points of Interest (PoI) data classified as 'commercial services' and includes locations of consultancies, employment and career agencies, hire service companies, advertising companies etc. Five per cent of trips were to work and four per cent of trips were from work. Figure 2.15 illustrates the distribution of trips among the remaining purpose categories - all of which account for three per cent of trips or less.

Figure 2.15: journeys by purpose from origin and purpose to destination



Base: all journeys excluding children playing in the street and trips transporting people or goods during working hours (11,090).

2.36 One quarter (25 per cent) of trips were missing either a 'to' or 'from' purpose code. Two per cent of trips had neither a 'to' or 'from' purpose. Excepting purpose categories 'home' and 'unknown', each remaining category consists of 800 trips or fewer, making further cross-tabulations unreliable.

Excluded journeys

Playing in the street

2.37 There are a number of journeys included in the GPS sourced data that would not be recorded using the paper travel diary. For example, the diary records time spent playing in the street, but does not consider this as travel activity.

2.38 If carried at all times, the GPS devices carried by 12-15 years old would have recorded any play in the street. Therefore, journeys for children aged 12-15 which combined walking, running or cycling, to or from home, friend's and family addresses and locations identified as 'personal other,' 'other' or 'unknown' were identified as 'playing in the street'. Some 176 journeys were classified this way, accounting for 3,551 miles or 7,748 minutes.

Transporting goods or people

2.39 Similarly, those working in jobs that involve the transportation of goods or people are not recorded in the NTS diary for practical purposes. This would include those moving or delivering goods (e.g. postmen), crew of specially-equipped vehicles (e.g. fire engines or ambulances), people paid to walk or cycle (e.g. police 'on the beat' or couriers) and many of the trips made by taxi drivers and driving instructors. The GPS pilot interview included a question which identified whether respondents fell into this category and then recorded their working hours for the travel week. Any trips which started during those hours can then be filtered from the data to ensure comparability with the existing traditional NTS. There were 177 journeys made during working hours by 40 people working in jobs which involved the transportation of goods or people. These accounted for 7,529 miles or 13,038 minutes.

Trips to/from work without fixed work sites

2.40 One of the challenges posed to the processing team, was how to identify trips to and from places of work when the respondent has no fixed workplace of work. TU/e did not incorporate this into the revised Trip Annotator; however, we can at least quantify how many trips are likely to have been affected by this.

2.41 The purpose of a trip was coded according to the personal information on addresses supplied during the pick-up interview, or PoI data associated to that location. Work addresses were one of the several addresses that respondents were asked to provide. Whilst it is easy enough to code trips

to work if a respondent works at a small number of fixed sites, it is not practicable to collect data for a large number of work places. Moreover, if someone does not have a fixed work site - consider a self employed plumber - it becomes increasingly difficult to code the purpose of their trip depending on the information held on the location. For example, work related trips made by a plumber to a grocery shop, dental surgery or restaurant would be coded as grocery shopping, personal medical and eating or drinking respectively. The GPS pilot included a question which identified whether respondents worked at many different sites. These people were then asked to provide their work hours, so that these trips could be identified.

2.42 There were 91 respondents who said they worked at different places. Of these, 59 had carried and returned a GPS device and did not work in jobs which included the transportation of goods or people. Eight of these could not recall or preferred not to provide their work hours for this week, another eight did not work during the travel week and 43 had worked. During the travel week these people made 209 trips during working hours. There are insufficient trips to tabulate these journeys by purpose; however, none were coded as being to or from work. Thirty-eight per cent of trips were coded as being from home (which is entirely feasible), 11 per cent were identified as return trips from locations for 'other personal' reasons, 10 per cent from 'sports' activities and 23 per cent from locations where the purpose of origin was 'unknown'. Similarly, 39 per cent of trips were coded as homebound (again, entirely feasible), 15 per cent were trips to locations for 'other personal' reasons and 20 per cent were to places where the purpose was 'unknown'.

3. Comparison to the diary data

Summary

3.1 This chapter compares the results from the processed personal travel data collected using GPS devices; to un-weighted results from the personal travel data collected using the main NTS survey during the same period. That is, data provided by the 1,726 respondents aged 12 and over, living at fully participating households drawn from the February and March 2011 NTS sample who completed a travel diary for travel weeks starting on or after 11 February 2011. This is **not** an early release of NTS 2011 data. Published NTS data are subject to complex processing, imputation and weighting prior to analysis. The NTS diary results presented here should therefore not be used in any other context.

3.2 Table 3.16 and Table 3.17 detail comparative statistics for the GPS and diary trip and stage data. The average number of stages per trip was 1.08 for the GPS data and 1.10 for the diary data. Trip rates based on the GPS data are lower than for the diary data: the GPS data estimates that each respondent makes an average 645 journeys per person per year[7] compared to an average of 934 journeys per person per year based on the diary data. Even if we compare the adjusted GPS trip rate in paragraph 2.33 that accounts for the apparent tail off in trips during the latter part of the travel week (774 trips per person per year) the GPS trip rate is still substantially lower than the diary trip rate.

3.3 However, the GPS trips were longer than the diary trips, in terms of mileage and duration: the average trip length in miles and minutes was 24 miles and 51 minutes for the GPS data and 6 miles and 21 minutes for the diary data. That is, the average GPS trip rate was 69 per cent of that for the diary, but the average trip mileage and duration were 370 per cent and 240 per cent of the diary estimates. If we compare the average journey time and distance per person, per year these differences aren't quite so pronounced. The average journey time per person, per year was 544 hours based on the GPS data and 328 hours based on the diary data (166 per cent), and the average journey distance per person, per

[7] As these estimates as based on only two months of data, these are not true annual rates, but equivalent estimates based on the available data.

year was 15, 429 miles based on the GPS data and 6,652 miles based on the diary data (232 per cent).

Table 3.16: GPS pilot and NTS diary stages (February- March 2011)		
	GPS	Diary
Sample size	897	1,726
Total number of stages	11,979	34,080
Total stage time (minutes)	567,065	665,691
Total stage distance (miles)	276,512	221,432
Average stage distance (miles)	23	6
Average stage time (minutes)	47	20
Stages per person per year	696	1,030
Average stage time per person per year (hours)	549	335
Average stage distance per person per year (miles)	16,073	6,689

Notes: GPS data excludes 'waiting' stages. Diary data for short walks collected on day seven are multiplied by a factor of seven; are based only on data from fully participating households, and exclude data for those aged under 12.

Table 3.17: GPS pilot and NTS diary journeys (February- March 2011)		
	GPS	Diary
Sample size	897	1,726
Total number of journeys	11,090	30,904
Total journey time (minutes)	561,114	652,160
Total journey distance (miles)	265,433	220,205
Average journey distance (miles)	24	6
Average journey time (minutes)	51	21
Journeys per person per year	645	934
Average journey time per person per year (hours)	544	328
Average journey distance per person per year (distance)	15,429	6,652

Notes: GPS data excludes journeys flagged 'children playing in the street' and trips made by those working in transportation jobs during working hours. Diary data for short walks collected on day seven are multiplied by a factor of seven; are based only on data from fully participating households, and exclude data for those aged under 12.

3.4 It is therefore reasonable to conclude that the trip estimates resulting from the GPS pilot are not similar to those from the diary. Nor are they consistently higher or lower than the diary results. When compared to the diary data, the GPS pilot underestimates the number of trips, yet overestimates the length of stages. This could be because some trips are missing; inaccuracies in the GPS data have distorted the length of journeys and/or some trips have been merged together. This could also explain the higher rates of trips to and from home in the GPS data trips if trips were merged into daily 'tours'. Previous research has also indicated that stray and inaccurate GPS points can easily add distance to journey traces. If some of these have been included in the Trace Annotator processing it is possible that they may be inflating the distance of trips.

3.5 Please note that as explained in paragraph 2.20 the NTS travel diary only collects data on short walks for day seven; therefore the number of short walk stages and journeys from the diary data has been weighted by a factor of seven for comparison to the GPS data.

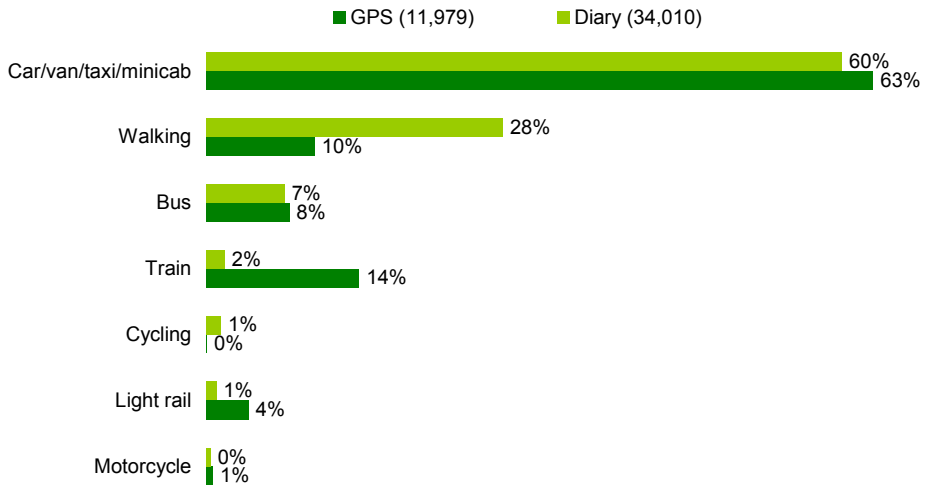
The distribution of GPS pilot and NTS diary stages and trips

Stages by mode of transport

3.6 Sixty-three per cent of GPS stages were allocated to 'car' mode. The travel diary identifies car, van or lorry, taxi and minicab as separate entities. Grouping these together into one category for comparison to the GPS indicates that a similar proportion of diary stages - 60 per cent - were made by 'car'. Ten per cent of GPS trips were inferred to be walking trips. This is less than half the 28 per cent of diary stages identified as walking trips. The proportion of stages made by bus according to the GPS and diary was similar (eight per cent and seven per cent respectively); however, the GPS processing inferred a far greater number of stages were made by train or light rail than in the diary. Fourteen per cent of GPS stages were allocated as being rail travel and four per cent were identified as light rail travel. In contrast just two per cent of diary stages were made by rail, and one per cent of stages were made using light rail.

3.7 It is difficult to comment on the number of trips allocated to cycling and motorcycling due to the relatively low number of trips made this way (just 474 and 135 in the diary data), however, in identifying just seven cycling stages and 75 motorcycle stages it would appear that the GPS processing hasn't managed these types of trips well, despite the apparent success of using the additional accelerometer recorded in the supplementary processing report. Figure 3.18 illustrates the distribution of GPS and diary stages by mode of transport.

Figure 3.18: stages by mode



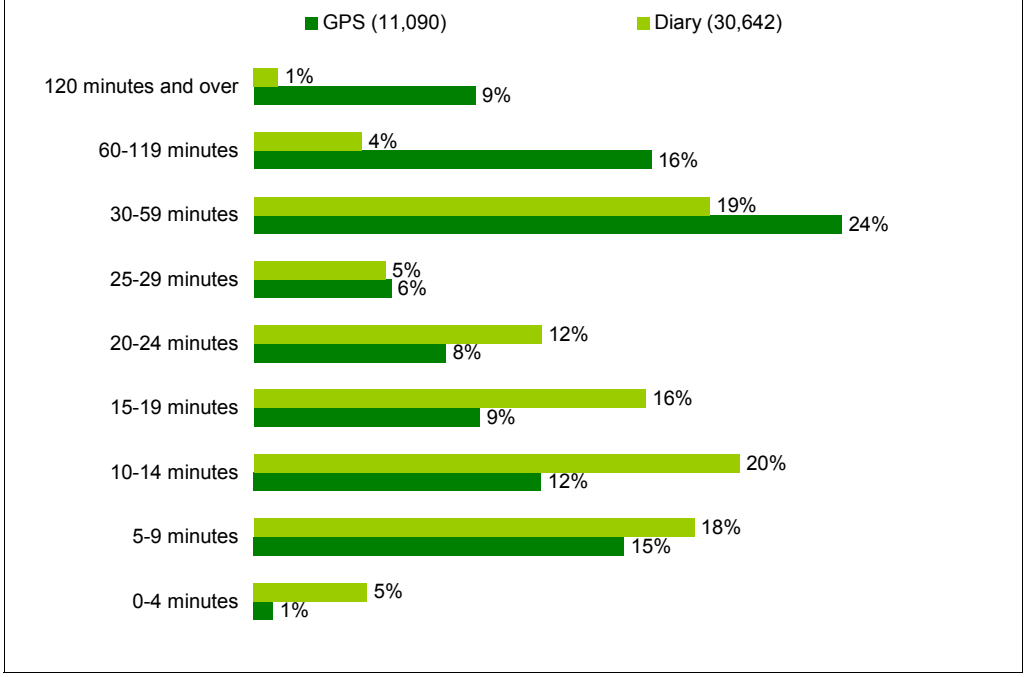
Base: shown in brackets

3.8 Just five per cent of GPS trips were short walks, whereas 16 per cent of diary trips were classified thus. The GPS methodology has therefore been less successful at identifying walking trips, particularly short walking trips. For practical reasons walking trips tend to be shorter than trips involving other modes, so we could anticipate the average length of GPS trips to be longer than the diary trips given that the GPS data contains fewer walking trips.

Length of journey

3.9 As observed in the summary statistics at the start of this chapter, the GPS trip rate was around two-thirds that of the diary data; however the average GPS trip length was approximately twice that of the average diary trip length and nearly four times the trip duration. This indicates that the GPS trips were typically longer than the diary trips and is further verified by Figure 3.19 which shows the distribution of time taken to complete journeys. The proportion of diary trips within each of the five minute bands between 0-24 minutes travel time is substantially higher than the proportion of GPS trips. In contrast, the proportion of GPS trips within each of the time bands with travel time of 30 or more is notably higher than the proportion of diary trips. In summary, 76 per cent of diary trips took less than 30 minutes to complete, whereas 49 per cent of GPS trips took at least 30 minutes to complete.

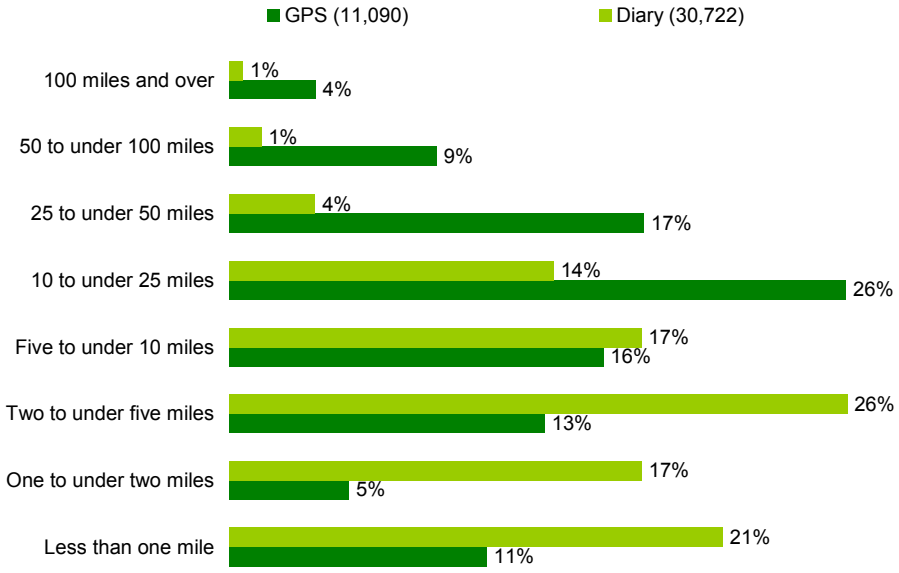
Figure 3.19: journey length (minutes)



Base: shown in brackets

3.10 Similarly, GPS trips tend to have longer mileage than the diary trips. Thirty-eight per cent of diary trips were less than two mile long and 64 per cent were less than five miles long compared to 16 per cent and 29 per cent of GPS trips respectively. Approximately half of the GPS trips (55 per cent) were at least 10 miles long compared to just 19 per cent of diary trips.

Figure 3.20: journey length (miles)

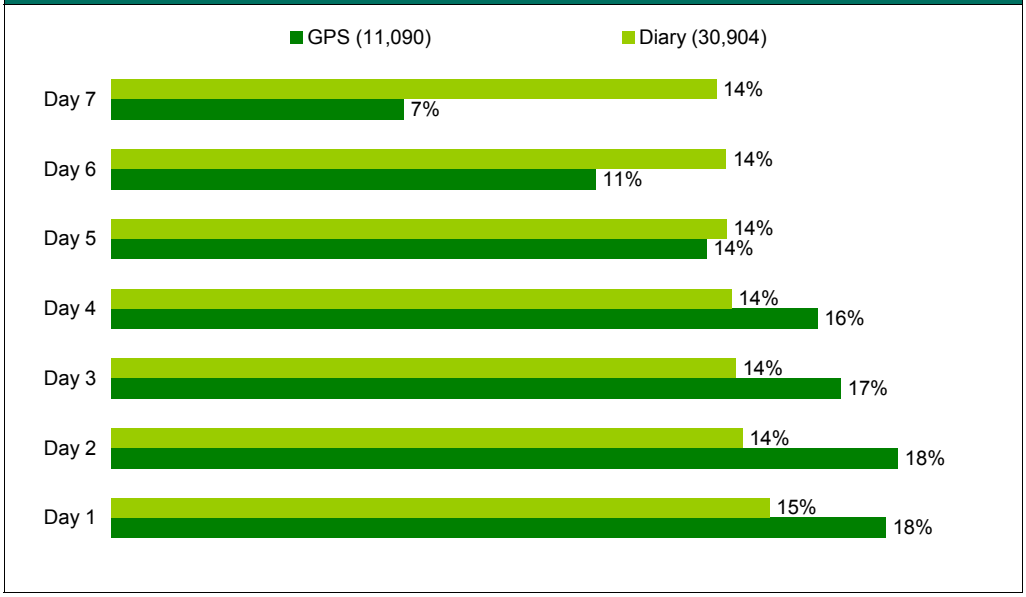


Base: shown in brackets

Journey by day of travel week

3.11 The diary series in Figure 3.21 demonstrates how the NTS travel week allocation methodology usually ensures that there is very little variation in the number of trips by travel day. Please note that it is to be anticipated that diary data variance is reduced marginally as the number of short walks on days 1-6 is exactly same as for day seven. However, while the diary data is very nearly uniformly distributed across the travel week travel days, with day one accounting for 15 per cent of trips and days 2-7 each accounting for 14 per cent of trips, this is not the case for the GPS data. Although the first four days of the travel week account for similar proportions of the GPS trips (16-18 per cent), the number of trips starts to tail off on days 5-7, falling to 14 per cent, 11 per cent and seven per cent.

Figure 3.21: journeys by day of travel week



Base: shown in brackets

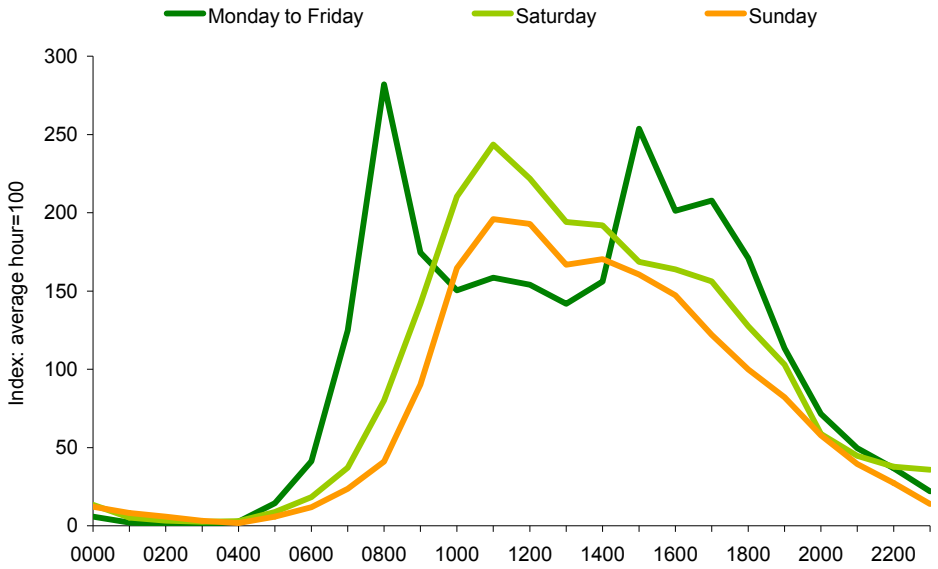
3.12 It is recognised that the reporting of trips in a weekly travel diary can tail off to some extent due to respondent disengagement or forgetfulness; however, theoretically this should not happen with GPS data collection methods. It could be that respondents forgot to charge or carry the devices with them during the latter part of the week, but Table 2.1 suggests that this is not the case. It is possible, however, that respondents misreported having not forgotten to charge or carry the device, or had genuinely forgotten that they had not carried or charged the device when asked. Another possibility is that there was some

unknown issue relating to the device and how it recorded data, although we have no evidence to suggest this was the case.

Journey by start time

3.13 Traditionally the NTS identifies distinct travel patterns on weekdays and weekends (Figure 3.22). On weekdays there are two peaks in travel coinciding with 'rush hour' activity: one between and 08.00 and 08.59 and another between 15.00 and 15.59 (mainly driven by education trips). At the weekend there is one peak between 11.00-11.59.

Figure 3.22 trips in progress by time of day and day of week 2010[8]

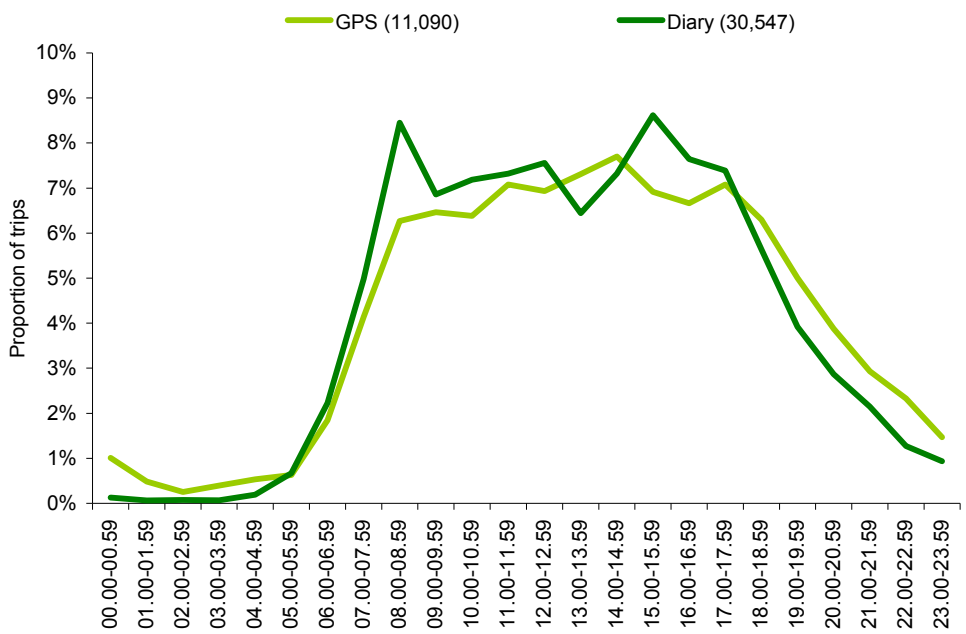


Source: GB National Travel Survey, 2010
Base: 251,000 trips Mon-Friday, 44,000 trips Saturday and 36,000 Sunday

[8] http://assets.dft.gov.uk/statistics/tables/nts0501.xls

3.14 Figure 3.23 presents the number of trips for the GPS pilot and the main NTS diary data for the same period by start time. It does not present the data separately for weekdays and weekdays as we only have a limited data set of two months rather than a full year's data, however, the two week day peaks between 08.00 and 08.59 and 15.00 and 15.59 are still apparent in the diary data. It is notable that these are not visible in the GPS trip data.

Figure 3.23: journeys by start time



Base: shown in brackets

Purpose of journey

3.15 Before it is possible to compare the results for GPS and diary trip data by purpose, we must harmonise the existing detailed diary purpose codes with the simpler GPS codes derived from personal address data and GIS PoI categories. Table 3.23 sets out how this has been done. These are not by any means perfect matches: if day trips are coded into 'holidays' then the code share 'just walk' must also be coded as 'holiday. Similarly, if 'escort - other shopping' is coded into 'shopping' then so is 'escort - other personal' as it also shares the same code.

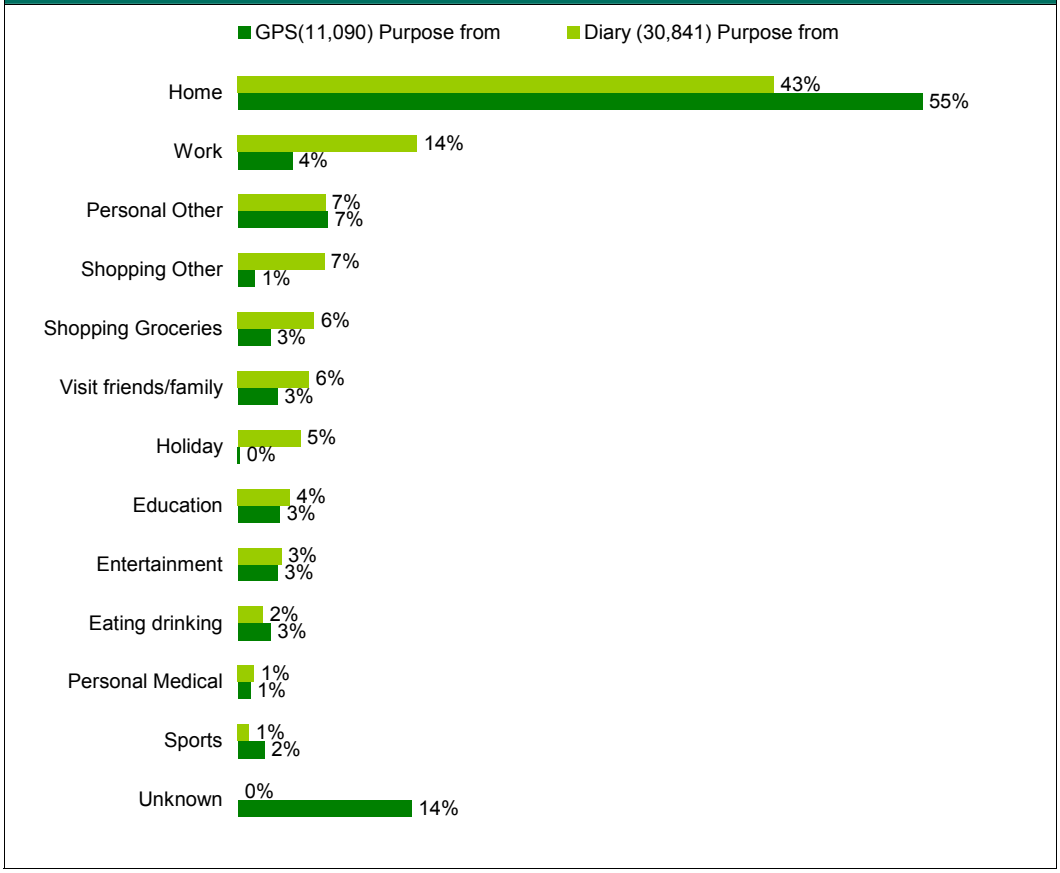
Table 3.23: matching the NTS diary and GPS pilot purpose code frames

GPS	Diary
Home	Home
Work	Work, In the course of work, Escort - work, Escort - in course of work
Education	Education, Escort - education
Shopping Groceries	Food and grocery shopping
Shopping Other	Other types of shopping , Escort - shopping/personal
Personal Medical	Personal business - medical
Personal Other	Personal business - other, Other escort, Escort - home (not own), Other non-escort
Eating drinking	Eat/drink - alone or at work, Eat/drink - other occasions
Friend	Visit friends or family
Entertainment	Entertainment/public activity, Other social
Sports	Sport (participate)
Holiday	Holiday base, Day Trip, Just walk
Unknown	-

3.16 Figure 3.24 and Figure 3.25 show the distribution of journeys by purpose from the origin and to the destination. The purpose of journey is one of the most important and unique features of the NTS to both DfT and external users. Therefore if GPS were to have a future as part of the NTS, it is crucial that the processing system can accurately infer the purpose of a journey. It is evident from these charts that there are three key areas of concern when comparing the GPS pilot results to the diary results with respect to the purpose of journey.

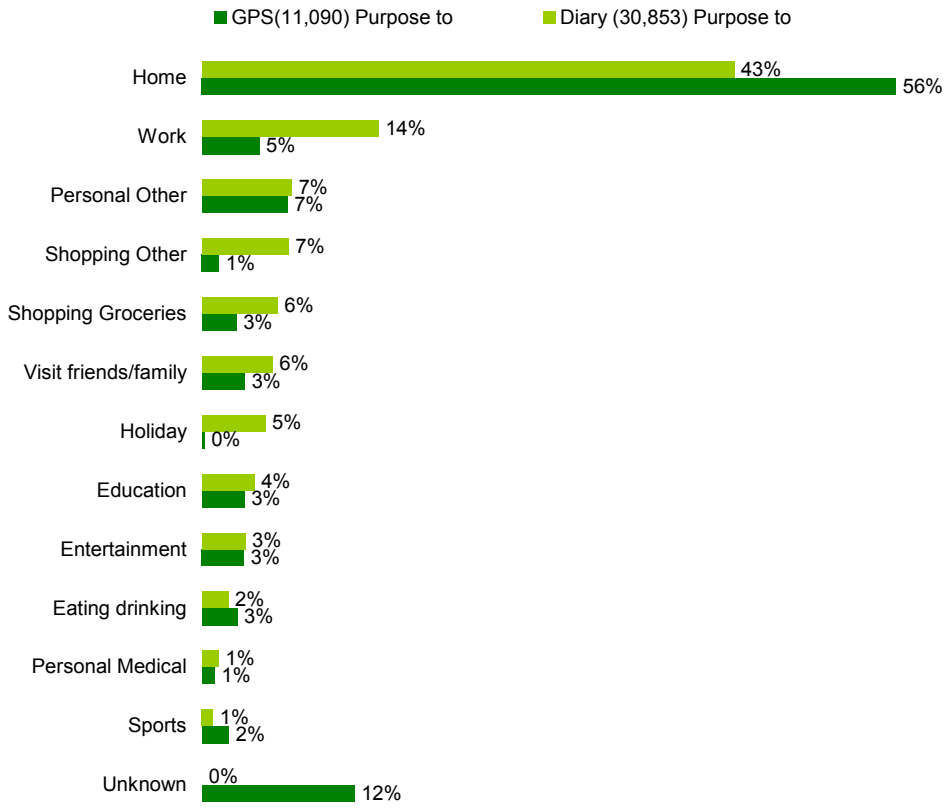
- The GPS data overestimates the number of journeys to and from home. Forty-three per cent of diary trips were from home and 43 per cent were to home while 55 per cent of GPS trips were coded from home and 56 per cent were coded to home.
- The GPS data underestimates the number of trips to and from work. Fourteen per cent of diary trips were trips from work and the same proportion were coded as trips to work. Just four per cent of GPS trips were coded 'to work' and five per cent 'from work'. The issue raised in paragraphs 2.40- 2.42 regarding the classification of trips to or from work as reliant upon one or a limited number of fixed work sites may explain some of the difference in these estimates, but not all of it.
- Fourteen per cent of trips 'from' a destination and 12 per cent 'to' a destination were coded as 'unknown'. This category was applied to sites where no information was available or if the information available was not able to distinguish between the purpose codes presented here (i.e. mixed used locations). Whilst we had always anticipated a small number of trips may fall into such a category under a GPS methodology, not knowing the purpose of one in seven trips is undesirable.

Figure 3.24: journey purpose by origin



Base: shown in brackets

Figure 3.25: journey purpose by destination



Base: shown in brackets

3.17 In addition to the address data usually collected in the NTS interview (respondent's usual place of work) respondents were asked to supply addresses for a wide range of locations that they may have visited during the travel week to supplement the GIS PoI information and code the purpose of visits to those addresses. These included:

- Children's nursery/school/college;
- Supermarkets;
- Gym or fitness centres;
- Cinemas or theatres;
- Other leisure facilities;
- Dentist's or Doctor's surgeries or hospitals;
- Work appointments;
- Community centres, churches or religious centres;
- Evening or weekend classes;

- Classes or clubs for children under the age of 12, and
- Friend's and family homes.

3.18 It was anticipated that many of these addresses would be incomplete, for example, a respondent might know that they use the Tesco supermarket on 'Smith Square, London' but they would be unlikely to know the street number or postcode. Therefore before the address data was supplied to TU/e for data processing, all addresses were run through an automated address matching program. Addresses that were still incomplete after this were subjected to manual look-up to ensure that addresses were as complete as possible.

3.19 It was hoped that using this personal data to assist in the coding of purpose of trip would reduce or eliminate the number of 'unknown' locations. TU/e estimated that for the journeys where purpose was successfully inferred, 73 per cent were coded using personal data and the remainder were coded using GIS PoI data. We do not have information on what proportion of trips coded using personal data could have been coded equally successfully using PoI data. It would be interesting to see such results for each purpose category, to see whether any categories could have been largely coded relying upon PoI data, rather than needing to collect personal address information for all purpose categories.

Estimating car occupancy and taxi usage using interview data

3.20 If the NTS were to switch to using a GPS methodology, there are a number of data items currently collected using the travel diary that are required by the department. At present the diary instructions ask respondents to:

- identify whether they were the driver or passenger during each of their car/van trips;
- supply the number of people travelling in a vehicle for each car or van journey, and
- record any taxis trips, the number of occupants and the value of the fare paid by the respondent.

3.21 Questions on these data items were added to the pick-up interview to investigate whether it is possible to collect robust data on car occupancy and taxi usage without the travel diary. Please note that where the term 'taxi' is used in this text, it also refers to minicabs.

3.22 The results to these questions and a 'best-fit' comparison to data from the travel diary data items are summarised in Table 3.26a and Table 3.26b. Some caution must be exercised when comparing these results as the diary data is mostly based on the number of stages recorded in the diary and the GPS pilot data is based on the number of respondents who provided a valid answer to the questions asked in the pick-up interview.

Table 3.26a : Car and taxi data based on diary trip data

		Base
Proportion of car trips respondent is driver	75%	19,367 stages
Mean occupancy of car	1.8	19,367 stages
Made taxi trip in week	12%	1,726 people
Average number of taxi trips per person	2.0	400 stages
Taxi fare	£6.09	400 stages
Mean occupancy of taxi	2.2	400 stages

Table 3.26b : Car and taxi data based on GPS pilot interview data

		Base: (people)
Proportion of car trips respondent is driver	59%	853*
Mean occupancy of car	2.3	868
Made taxi trip in week	9%	912
Average number of taxi trips	1.9	84
Taxi fare	£8.39	84
Mean occupancy of taxi	2.1	84

*Excludes five per cent 'can't remember'.

3.23 The NTS travel diary recorded that respondents had been the driver of the car in 75 per cent of car stages. However, 59 per cent of GPS pilot respondents said that they had been the driver of the car in their most recent car trip. This suggests that a question about the most recent car trip is not a suitable alternative to diary based estimates on the proportion of trips was a driver or passenger.

3.24 The diary estimated car occupancy to be 1.8 persons and the question placed on the GPS pilot questionnaire regarding the most recent car trip the respondent had made estimated this to be 2.3 persons. This result indicates that the interview question is not a suitable proxy for the diary measurement of car occupancy. It is possible that the question does not work well as the 'most recent' car trip is quite likely to have been a return journey home and therefore unrepresentative of all car trips.

3.25 While the diary recorded that 12 per cent of respondents had used a taxi during the travel week, nine per cent of GPS pilot respondents said they had made at least one trip using a taxi. Diary and GPS pilot respondents who had used a taxi during the travel week, had typically made two trips by taxi during the travel week. Please note that the base for this estimate is small (just 84 respondents or 400 stages) and should be considered indicative rather than representative.

3.26 The mean amount paid towards the fare of a taxi by each respondent was £6.09 according to the travel diary and £8.39 according to the GPS pilot interview, however; owing to the small sample sizes, these estimates are not significantly different[9]. The two sources provided similar estimates of taxi occupancy: 2.2 (diary) and 2.1 (GPS pilot), however, bearing in mind the sample sizes, this could be largely coincidental.

[9] Tested for statistical significance at the 95 per cent level of confidence.

4. Conclusions

4.1 In the main NTS one of the key activities of the interviewer during the pick-up interview is to collect travel diaries for each member of the household; check that these have been fully completed and that there are no incomplete or unclear entries, and query and correct any such entries with the members of the household who are present. If a respondent forgets to record a trip or trips in the diary on a particular day, they can always enter this later in the week or during the pick-up interview. In using an automated system such as a GPS device, there are a number of ways in which trip data could be missing which can't be corrected retrospectively. The NTS GPS pilot found that on average, seven per cent of respondents made journeys but forgot to carry the device with them for some or all of the trips made on each day of the travel week. We are unable to estimate what proportion of trips were missing due to devices running out of battery power, but three per cent of respondents reported having forgotten to charge their GPS device each day of the travel week.

4.2 There were far fewer trips and stages in the GPS data (645 trips per person year compared to 934 trips per person in the diary data) but the time taken to complete these was much longer (average journey time per person per year was 544 hours compared to 328 hours based on the diary data) and the average journey length was also much longer (15,429 miles per person per year, compared to 6,652 miles according to the diary).

4.3 Other noticeable discrepancies between GPS and diary data include:

- The GPS data identified far fewer walking trips than the diary (10 per cent compared to 28 per cent) and approximately five per cent of GPS trips were short walks, compared to 16 per cent of diary trips. In contrast the GPS data identified seven times more rail trips than the diary (14 per cent compared to two per cent);
- The number of trips per travel day is evenly distributed across the NTS diary travel weeks; however the number of GPS trips is similar for days 1-4 (16-18 per cent) then falls to 14 per cent, 11 per cent and seven per cent on days 5-7. This could be due to respondents forgetting to charge or carry the devices as the travel week progresses, despite their reporting this was not the case, and

- NTS trips typically peak during the morning and afternoon rush hours, (08.00-08.59 and 15.00-15.59) however; the GPS trip does not contain these travel patterns and instead peaks between 14.00-14.59.

4.4 One of the most important factors of assessing the suitability of GPS data collection as a practicable method of data collection for the NTS is whether GPS processing can infer the purpose of a trip within acceptable parameters. The GPS data identified that 55-56 per cent of trips were to or from home compared to 43 per cent of diary trips. It also estimated that 4-5 per cent of journeys were to/from work compared to 14 per cent based on the diary data. Furthermore, some 14 per cent of trips were coded as purpose 'unknown'. One quarter (25 per cent) of trips were missing either a 'to' or 'from' purpose code and two per cent of trips had neither a 'to' or 'from' purpose. There are also a number of substantial differences in the proportion of trips coded purpose to/from shopping trips, visiting friends or family and holidays. Some of these differences may be accounted for by the difference between the code frames, however, these are less likely to affect categories such as home, work or the homes of friend's and family.

4.5 There are a number of issues that could be further investigated. For example, 73 per cent of journeys for which the purpose was inferred were coded personal address data. The remainder were coded using PoI data from the Ordnance Survey database. However, we do not know whether there were any particular address types where categories of purpose could be coded successfully without the need for collecting personal data. The data collected on hours of work during the travel week for those who transport goods or people for a living and those who work at no fixed work place was not utilised in the processing of GPS data. It is possible that more could have been done with this to estimate the number of 'eligible' NTS trips and improve the prediction of trip purpose.

4.6 It would appear that asking questions in the pick-up interview about whether the respondent was the driver or a passenger in their last car trip and how many people were in the car is unlikely to provide a comparable estimate to the stage data currently collected on these issues by the travel diary. It is difficult to assess whether the use of questions on the use of taxis provides a robust alternative to the diary data on taxi usage owing to small sample sizes, however, based on what data we did collect (400 stages and 84 respondents) there is no evidence to suggest that the two are not compatible.

4.7 On the basis of the numerous and substantial differences between the results for the GPS pilot survey and the NTS diary data for the same period, we conclude that GPS devices do not provide an acceptable

alternative data collection tool when used in the context of the NTS methodology and the Trace Annotator processing system.