Lowe *et al.*
European Transport Research Review (2024) 16:45
https://doi.org/10.1186/s12544-024-00666-w

European Transport Research Review

ORIGINAL PAPER

Personalizing travel behaviour change interventions using the trans-theoretical model and multimodality data

Warnakulasooriya Umesh Ashen Lowe[1*], Leonhard Lades[2] and Páraic Carroll[3]

Abstract

Introduction Behaviourally informed soft policies, such as nudges, have become popular in areas like health, environment, and energy use as cost-effective instruments to change behaviour and decision-making. However, the effectiveness of soft policies in the transport sector is modest at best. One reason for this relative ineffectiveness might be their one-size-fits-all nature, and personalizing soft interventions has been suggested to increase their effectiveness. The Trans-theoretical Model (TTM) suggests that people progress through five stages of behavioural change, from pre-contemplating a behaviour to maintaining the behaviour, and behavioural interventions could be designed for specific stages. However, it is not always feasible to conduct surveys to place people at different stages of the TTM.

Methods This paper explores whether it is possible to use multimodality data taken from a travel diary to place people at different stages of the TTM. The analysis uses an existing dataset from 826 respondents that includes self-reported TTM stages regarding cycling and data on multimodality. In the analysis, the multimodality data are used to allocate respondents to categories and assign them to TTM stages. The performances of the stage assignment approaches are evaluated using the self-reported TTM data and confusion matrices.

Findings The accuracy of the allocation of participants to TTM stages using multimodality data is approximately 75%. The accuracy is higher for early stages (pre-contemplation) and later stages (maintenance) of the TTM. A data-driven approach to dealing with multimodality data performs slightly better than an approach that relies on pre-defined categorization.

Conclusion The paper suggests that it will be possible in the future to personalise behavioural interventions according to the stages of the TTM even in the absence of self-reported survey data that classifies people to TTM stages if objective multimodality data are available.

Keywords Travel behaviour interventions, Multimodality measurements, Stage model, Behavioural change

1 Introduction

Significant transformations in travel behaviour are necessary to change existing mobility patterns that are at odds with achieving environmental targets. Governments all over the world employ a range of travel behaviour intervention strategies, encompassing financial incentives and disincentives, infrastructure enhancements, as well as marketing, and information initiatives [40]. These policy measures can be broadly categorized into two distinct groups: "hard" interventions seek to reshape social contexts and structures, and "soft" interventions aim to

*Correspondence:
Warnakulasooriya Umesh Ashen Lowe
warnakulasooriya.lowe@ucdconnect.ie
[1] School of Civil Engineering, University College Dublin, Dublin, Ireland
[2] School of Architecture, Planning and Environmental Policy, University College Dublin, Dublin, Ireland
[3] Faculty of Architecture Building and Planning, The University of Melbourne, Glyn Davis Building, Parkville Campus, Australia

influence individuals' perceptions, beliefs, attitudes, values, and norms [39] and influence behaviour through nudges that change how choices are presented to people [22].

Soft interventions have gained significant attention in recent policy frameworks. The "Shift" measures outlined in Ireland's 2023 Climate Action Plan, which are mostly soft in nature, aim to promote a modal shift towards more sustainable modes of transportation. The report emphasizes the importance of complementing infrastructure improvements with effective behavioural measures [11]. According to the Department of Health in Scotland, the provision of personalised travel plans can successfully achieve a modal shift towards sustainable travel modes [49]. The "Green Lane project" in the London Underground is another example of using soft interventions to achieve behaviour change [31]. Additionally, the introduction of green license plates in several jurisdictions, such as Canada, Hungary, China, and Norway, serves as an effort to encourage the use of environmentally friendly vehicles [45]. City-wide cycling campaigns such as "Cycling May" [7] are becoming increasingly popular tools for achieving significant travel mode shifts.

While these soft interventions can be politically more feasible, their impact on travel behaviour change is limited (e.g., [23]). A recent meta-analysis spanning three decades of soft interventions revealed a modest impact of a 7% reduction in car usage following the implementation of soft interventions [39]. One potential reason for the limited impact of soft interventions to change travel behaviour is their generic, one-size-fits-all nature which disregards that people's responses to interventions can vary and overlooks the differences in how different people react to different treatments [43]. For instance, Tang et al. [46] highlighted the limitations of some one-size-fits-all strategies used in many cities in China to address helmet use among e-bikers. Personalizing the delivery of soft interventions can resolve the issue by determining the suitability of different types of soft interventions for different groups of people [27, 34]. A promising way to distinguish between these groups is to segment the target population depending on how motivated they are to change their behaviour. For example, soft interventions that trigger some re-thinking in people who have just begun thinking about travelling more sustainably might be entirely ineffective for people who have already formed intentions to change the way they travel and are looking into converting this intention into action in the near future [32].

A popular stage model of behavioural change is the transtheoretical model (TTM) which was suggested by Prochaska and Velicer [35]. The model divides people into various stages depending on how advanced they are in the process of behavioural change (see Fig. 1). The model has been used to describe stages of behavioural change in domains such as diet [41], physical activity [2], adolescent smoking [18], dental health [29], and travel behaviour modification [15]. The TTM typically consists of five stages: pre-contemplation, contemplation, preparation, action, and maintenance. Individuals in the pre-contemplation stage are unaware of problematic travel behaviour (e.g., single occupancy car journeys) and have no intention to change [17]. Individuals in the contemplation stage are aware of the problematic travel behaviour and are thinking about changing their behaviour. Individuals in the preparation stage intend to change their behaviour soon. Action stage individuals already use sustainable transport modes but not very often. Finally, individuals in the maintenance stage have sustained the use of sustainable transport modes for a while and try not to relapse.

Several studies have applied the TTM to examine behavioural change in the context of active travel. For example, Forward [14] provides empirical support for the

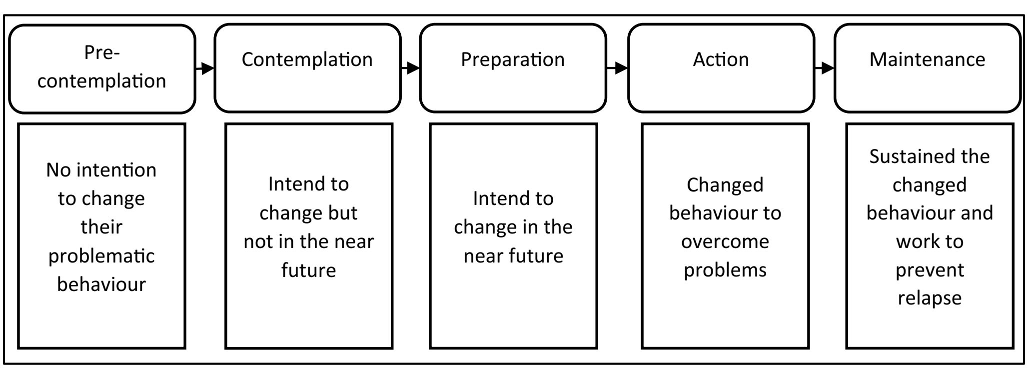

Fig. 1 Transtheoretical model stages and characteristics of individuals in each stage [48]

stage-based model and shows that the combination of psychological mechanisms (relationship between behavioural intention, attitudes, subjective norms, and perceived behavioural control) changed from stage to stage. Olsson et al. [33] found that personal norms, attitudes, and perceived behaviour control significantly contribute to explaining the differences between behavioural stages. Similarly, Thigpen et al. [47] emphasized the importance of travel-related attitudes in differentiating stages of behavioural change. Moreover, a few studies have utilized shifts between behavioural stages to evaluate behavioural change interventions [1, 16, 30, 32]. Understanding how ready people are to change their behaviour can lead to more effective interventions [48]. For instance, Sunio et al. [42] develop a behavioural intervention based on a self-regulated behaviour change model which allows users to self asses their respective stages and then get a stage-tailored diagnostic report which then produces a menu of recommendations. However, it has been proposed that different interventions should explicitly target the different stages proposed by the models [4, 33] rather than allowing individuals to select the intervention element. For instance, Mundorf et al. [30] assessed a video pilot intervention of sustainable travel options and health benefits, primarily targeted at pre-contemplators and contemplators. They found that it is effective in moving respondents towards increased readiness for sustainable travel change. Olsson et al. [32] observed a higher progression of individuals in the pre-action stage (which corresponds to the preparation stage in the TTM) due to a cycling trial intervention. Biehl et al. [6] reported that interventions which aim to enhance the ease and convenience of using sustainable travel may prove advantageous for individuals in the maintenance stage.

The typical way to determine how far advanced people are in their stages of travel behaviour change is to ask them in surveys. In some studies, respondents are presented with statements reflecting each behaviour change stage and are directly asked to indicate which statement is most applicable to them [14, 26, 33, 38]. Other studies assign people to a stage based on their responses to questions that do not directly refer to the stages of the TTM [5, 6, 17]. For instance, Biehl et al. [5] use a two-step verification process. In the first step, respondents are asked to choose a statement that best reflects their travel behaviour regarding a specific travel mode (statement 1: "I have never contemplated making a routine trip using this mode", statement 2 "I have contemplated making a routine trip using this mode", and statement 3 "I use this mode for at least one routine trip"). Based on the answer, in the second step, a follow-up question is asked to determine their stage of change (the follow-up question to statement 1 asks participants to indicate whether using the mode as a primary means of travel is a realistic alternative,the follow-up question to statement 2 asks participants to indicate whether they expect to use the mode as a primary means of travel in the near future; the follow-up question to statement 3 asks participants to indicate for how long they have been using the mode for a routine trip).

However, the use of the TTM to inform soft travel behaviour interventions is not widespread. One reason for this might be that determining how far advanced people are in their stages of travel behaviour change relies on self-reports in surveys which are rarely available to policymakers and have small sampling rates [51]. While using surveys to determine how far advanced people are in their stages of travel behaviour change is arguably the optimal approach, it is not always feasible to run survey studies with the population whose behaviour should be changed. Some studies have explored alternative methods that are based on objective measures [8, 19] in the context of exercise behaviour.

An alternative method to segment people into groups is to rely on objectively observable data about personal travel patterns and the different travel modes people use (i.e., data on "multimodality" which describes how people use more than one transport mode in a given period). The benefit of relying on multimodality data over relying on self-reported survey data is that multimodality data is often readily available to national transport authorities and that it is less invasive to obtain. Kroesen [24] shows that individuals who rely on multiple modes are more inclined to change their behaviour profiles over time compared to those who rely solely on one mode. According to Heinen [20], the higher the multimodality of an individual is, the higher the likelihood of reducing car use. Further, Heinen and Ogilvie [21] find that commuters with a higher level of baseline variability are more inclined to increase their active mode share while decreasing their car mode share. These findings suggest that multimodality may be an indicator of an increased likelihood to change behaviour towards more healthy and sustainable travel alternatives. As a result, it might be possible to identify groups of people who are more willing to change their travel behaviour using multimodality data.

This paper explores whether it is possible to use commonly available multimodality data to allocate people to stages of behavioural change based on the Trans-theoretical Model (TTM). This would allow the targeted design of soft travel policies to specific groups of travellers as suggested by the TTM using multimodality data that is frequently available to policymakers, hence boosting the effectiveness of soft behavioural change interventions. We rely on an existing dataset [5], because it is one of the few data sources that contains self-reported survey data indicating how far advanced study participants are in their TTM stages of travel behaviour change from driving to cycling as well as data on people's multimodality. We test whether different methodologies of clustering people using the multimodality data lead to clusters that map onto the five stages of the TTM survey data.

The remainder of the paper is organised as follows. The next section presents the data source, the pre-defined and data-driven approaches to multimodality measurements, and the TTM stage assignments based on the multimodality data. The results of this study are presented in the third section and discussed in the fourth section. The last section presents the conclusions and offers directions for future research.

2 Methodology

2.1 Data source

To map multimodality data onto the stages of the TTM, we rely on an existing travel diary dataset that was originally collected by Biehl et al. [5]. This is one of the few datasets that include data on both multimodality from a travel diary and self-reported information about placement in the TTM for travel behaviour change. The self-reported information on TTM stages is available for cycle use and walking in the dataset by Biel et al. (2018), and we focus on cycling in this paper. This survey was conducted over a 3-week period in February 2017 in six midwestern states (Indiana, Illinois, Michigan, Minnesota, Ohio, and Wisconsin) in the USA collecting travel information on three different trip purposes: work/school trips, shopping trips, and leisure trips. The present study considers all three types of trip purposes and utilizes an already screened and cleaned dataset which was shared with us by the original authors, consisting of data from 826 responses from the original sample. See Appendix A for summary statistics of the sample. The sample is composed of 58% females, the average age of participants is 37, 70% of participants have a college degree, 93% have a valid driving license, 79% own a bicycle, and 66% are employed full-time. Participants took an average of 21.85 trips, mostly for work/school (37.39%), leisure (35.70%), and shopping (26.91%). Across all trip types, private car use was the highest (46.16%), followed by walking (20.12%), ridesharing (16.13%), cycling (9.93%), and public transport (7.66%). For leisure trips also, private car use remained the highest (30.27%), with walking (27.17%), ridesharing (21.29%), cycling (14.43%), and public transport (6.84%). The shopping trips showed an even higher preference for private car use (52.31%), followed by ridesharing (19.54%), walking (17.32%), cycling (6.21%), and public transport (4.61%). For work/school trips private car use share peaked (62.24%), with walking (11.12%), public transport (10.16%), ridesharing (9.59%), and cycling (6.90%).

2.2 Multimodality measurements

Multimodality data can be used to classify travellers into different groups in at least two ways. First, researchers can classify people using pre-defined groups as multimodal or unimodal travellers based on their main method of transportation and whether they utilize other/specific modes. This approach does not consider the intensity with which people use these modes (see [50]). Second, researchers can use the data-driven approach which builds on unsupervised classification methods (see [20, 24]). For example, Heinen [20] used K-mean clustering on variables of share of trips made entirely and partially by a certain travel mode and identified three clusters (mainly unimodal car, mostly unimodal bicycle, and mostly multimodal). Kroesen [24] used latent class analysis on the share of trips by car, bicycle, and public transport and identified five clusters (strict bicycle user, strict car user, light traveller, joint car and bicycle user, and public transport users). In addition, continuous indicators can be used to quantify multimodality levels for each individual [3].

There are multiple indicators available to measure the level of multimodality such as the Herfindahl–Hirschman index (HHI) [36, 37], the objective mobility personal index (OM_PI) [12], the Gini index [10], the Theil index [13], the Dalton index [9], and the Atkinson index [13]. After conducting theoretical investigations and empirical experiments, Diana and Pirra [13] concluded that no single measure of multimodality consistently outperforms all the others in any circumstance. Furthermore, the precise choice of measure to use may not be critical as Heinen and Ogilvie [21] showed in their results that there are only small differences between the different indicators. Among these measures, HHI is regularly used as a measure of concentration [44] and is easy to calculate using travel diary data sets.

2.3 Pre-defined approach to multimodality measurements

We can use the pre-defined approach to segment people into groups based on their multimodality data. This approach uses the proportion of trips taken by the car and the bike. In line with Vij et al. [50], individuals who make less than 10% of their trips by car are classified as green multi-modal users (MM_{green}). Quasi multi-modal users (QMM_{car}) make between 10 and 90% of their trips by car. Quasi-uni-modal users (QUM_{car}) are those who drive 90% or more of their trips. Quasi multi-modal users and green multi-modal users are further subdivided into groups based on their shares of cycle trips. These subgroups include multi-modal users with 0% cycling share, multi-modal users with cycling share of 0% to 20%, and multi-modal users with a cycling share of 20% or more. We consider the threshold value of 20% as in Vij et al. [50]. For example, if an individual commutes to work by bicycle once every week and does not travel otherwise, their bicycle share would be once every 5 days (once every week), which equates to 20%. Hence, an individual with a bicycle share equal to or higher than 20% indicates that they have taken at least one trip per week by bicycle. Using this threshold value, the pre-defined approach identifies a total of 7 groups of multi-modal users. This is summarised in Table 1.

2.4 Data-driven approach to multimodality measurements

For the data-driven approach, we use k-means clustering to classify people using the multimodality data and map them onto the five stages of the TTM. We use four variables for clustering as summarised in Table 2. The Herfindahl–Hirschman index (HHI) is the sum of the squared values of the share of each mode within all commuting trips as shown in Eq. 1 [36]. It measures the concentration and emphasizes the importance of modes with large shares [44]. A normalized index ranges from 0 to 1, where the closer to 1 the more one mode dominates. For each individual i, the normalised HHI values are computed using Eq. 2 [20].

$$HHI_{basic} = \sum_{i=1}^{N} S_i^2 \tag{1}$$

$$HHI = \frac{HHI_{basic} - \frac{1}{N}}{\left(1 - \frac{1}{N}\right)} \tag{2}$$

In Eqs. 1 and 2, N represents the number of travel modes and S denotes the share of trips taken by a particular mode.

Table 2 Variables used for clustering

Variable	Description	Mean	S.D.
HHI	Herfindahl–Hirschman index	0.37	0.26
Count	Number of different modes used	3.23	1.13
Cycle share	The proportion of trips made by cycling	9.93	13.56
Rank	Priority given to cycling	2.96	1.03

Heinen [20] used “Count” which represents the number of modes used for trips. It ranges from 1 to 5 since five modes of transport (private car, rideshare, walking, cycling, and public transport) are considered in the present study. Just as the Herfindahl–Hirschman index, Count also depicts the multimodality characteristics. The third variable is the share of trips taken by bike. The fourth variable is the rank of cycling amongst all modes which indicates the priority given to cycling, with a rank of 1 if the bike is the most frequently used travel mode and a rank of 5 if the bike is the least frequently used mode. All four variables were considered for the cluster analysis using the k-mean algorithm. Wu and Kumar [52] state that k-means is one of the most widely used algorithms and it is relatively simple to implement and is fast when dividing a data set into the number of clusters as specified by the user. Table 2 also presents the means and standard deviations of the four measures in our dataset.

2.5 Allocating people to TTM stages using pre-defined multimodality groups

In order to allocate pre-defined groups (see Sect. 2.3) into TTM stages for cycle use, we consider the share of bicycle trips of each individual. We map the 7 groups presented in Table 1 (i.e., quasi-uni-modal users, green multi-modal users with zero cycle share, green multi-modal users with 20% or less cycle share, green multi-modal users with 20% or higher cycle share, quasi multi-modal users with

Table 1 Different pre-defined groups used in the present study

Group	Description	Count
QUM_{car}	A person who drove more than 90% of all trips with a car	72
MM_{green} ($S_{cy} \geq 20\%$)	A person who drove less than 10% of all trips with a car and cycled 20% or more of all trips	43
MM_{green} ($S_{cy} < 20\%$)	A person who drove less than 10% of all trips with a car, cycled more than 0% and less than 20% of all trips	39
MM_{green} ($S_{cy} = 0\%$)	A person who drove less than 10% of all trips with a car and did not cycle at all	54
QMM_{car} ($S_{cy} \geq 20\%$)	A person who drove between 10 and 90% of all trips and cycled more than 20%	118
QMM_{car} ($S_{cy} < 20\%$)	A person who drove between 10 and 90% of all trips with a car and cycled more than 0% and less than 20%	244
QMM_{car} ($S_{cy} = 0\%$)	A person who drove between 10 and 90% of all trips with a car and did not cycle at all	256

QUM_{car}, Quasi uni-modal user; QMM_{car}, Quasi multi-modal user; MM_{green}, Green multi-modal user; and S_{cy}, share of trips made by cycling

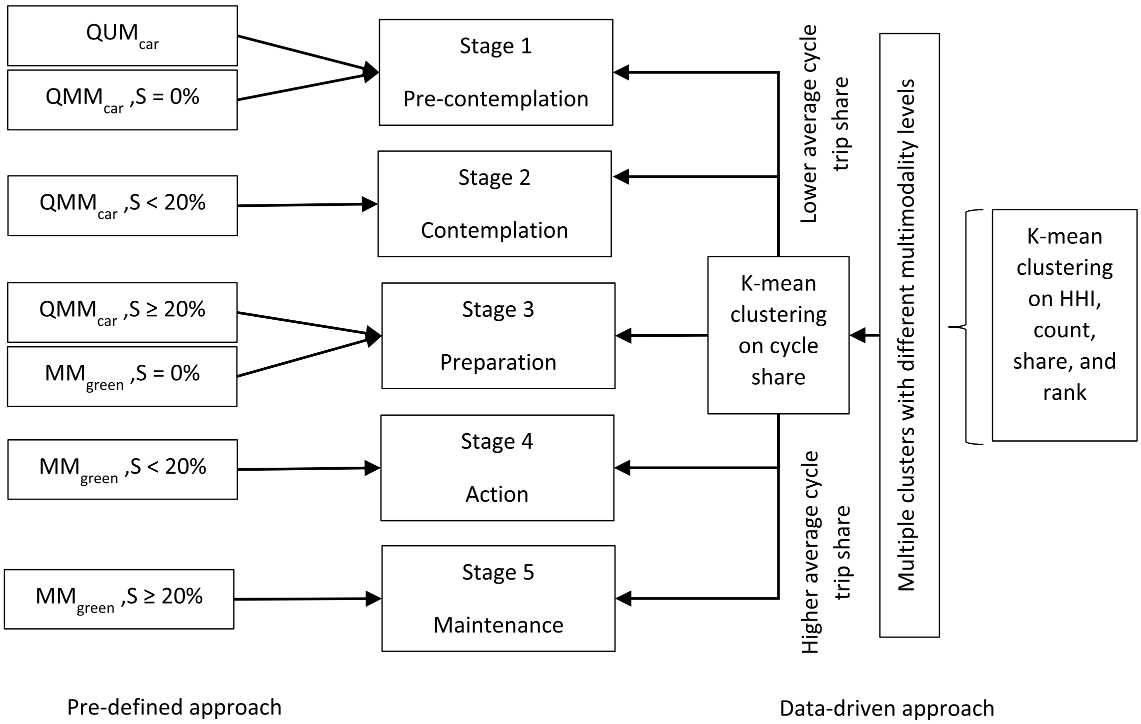

Fig. 2 TTM stage assignment process based on multimodality data using the pre-defined approach (on the left) and the data-driven approach (on the right). *Note* QUM_{car} = People who made 90% or above share of trips by car driving; QMM_{car} = People who made more than 10% (but less than 90%) of all trips by car driving; MM_{green} = People who made less than 10% of all trips by car driving; S = share of trips made by cycling

zero cycle share, quasi-multi-modal users with 20% or less cycle share and quasi multi-modal users with 20% or higher cycle share) to the 5 stages of the TTM presented in Fig. 1 (pre-contemplation, contemplation, preparation, action and maintenance) as illustrated in the left side of Fig. 2. We assign quasi-uni-modal users who heavily rely on driving to the pre-contemplation stage of the TTM, assuming that they do not consider the change to be possible. We also assign quasi-multi-modal users with zero-cycle trips to the pre-contemplation stage, as they drive often, and do not cycle at all. We assign quasi-multi-modal car users with a share of bicycle trips less than 20% to the contemplation stage where individuals have contemplated cycling but may not cycle as much as individuals in the preparation stage. We assign quasi multi-modal users with a share of cycle trips higher than 20% of the total trips to the preparation stage as these individuals take at least one trip per week by bicycle indicating a potential willingness to change their travel mode to cycling in the near future. Green multi-modal users drive less frequently, suggesting their inclination towards sustainable travel modes. Therefore, we assign green multi-modal users with a share of cycle trips of zero to the preparation stage. Given that their car use is minimal, and that they are multi-modal users, it is reasonable to assign them to this stage. We assigned the individuals with a share of cycle trips of 20% or greater among the green multi-modal users, to the maintenance stage, where individuals frequently use bicycles. We assign the remaining green multi-modal users to the action stage, where individuals cycle often but not as frequently as those in the maintenance stage.

2.6 Allocating people to TTM stages using data-driven multimodality groups

For the data-driven approach, we use the share of bicycle trips to assign the group of clusters to TTM stages. First, we identify the clusters by using the k-mean algorithm on HHI, rank, count, and bicycle share so that each cluster has average values for each variable. Then we use the k-means algorithm on the average cycle share of the cluster variable to derive 5 clusters from those clusters derived at first. For example, k-means clustering identifies 10 clusters from the dataset. Subsequently, we reapply the k-means algorithm to these 10 clusters to further derive 5 clusters. Then we assign the cluster with the highest average cycle trip share to stage 5 (the maintenance stage) and the one with the lowest average share of cycle trips to stage 1 (the pre-contemplation stage). This entire process is graphically depicted on the right side of Fig. 2.

2.7 The evaluation criteria

Following the assignment of individuals into TTM stages, we use confusion matrices to evaluate how accurately the clustering based on the multimodality data maps onto the self-reported classification into TTM stages. This evaluation method is built upon four criteria. The true positives (TP) are the number of correctly included individuals, the true negatives (TN) are the number of correctly excluded individuals, the false positives (FP) are the number of wrongly included individuals, and the false negatives (FN) are the number of wrongly excluded individuals. These four criteria are used to estimate the accuracy, precision, recall, and the "F1-score" using the following equations as suggested by Mirzahossein et al. [28]:

$$Total\ observations = (TP + FP + TN + FN) \tag{3}$$

$$Accuracy = \frac{(TP + TN)}{Total\ observations} \tag{4}$$

$$Precision = \frac{TP}{(TP + FP)} \tag{5}$$

$$Recall = \frac{TP}{(TP + FN)} \tag{6}$$

$$F1 - score = 2 \times \frac{Precision \times Recall}{(Precision + Recall)} \tag{7}$$

The accuracy indicates how many true stage allocations are made. The precision gives the proportion of true allocations made to the total allocations made. The recall indicates the proportion of true allocations made to the total allocations that should be true. The F1-score attempts to balance precision and recall by combining them which makes it more appropriate for situations where the class distribution is uneven [28]. Alongside the above four measures, Cohen's kappa coefficient (k) was estimated using Eqs. (8) to (10). Cohen's kappa coefficient also offers a reliable evaluation in situations where the classes are not evenly distributed [28], which can pose challenges for evaluating the performance of classification approaches.

$$p_o(Observational\ probability\ of\ agreement) = \frac{TP + TN}{Total\ observations} \tag{8}$$

$$p_e(Expected\ probability\ of\ agreement) = \frac{TP + FP}{Total\ observations} \times \frac{TP + FN}{Total\ observations} + \frac{FN + TN}{Total\ observations} \times \frac{FP + TN}{Total\ observations} \tag{9}$$

$$k = \frac{p_o - p_e}{1 - p_e} \tag{10}$$

Cohen's kappa is used to measure the agreement between classified data [25]. The coefficient varies from 0 to 1 at six levels where larger values indicate higher efficiency and effectiveness of the classification approach. Table 3 presents the interpretation of the different kappa values as suggested by Mirzahossein et al. [28].

Table 3 Kappa statistic measures for categorical data. Source: [28]

Kappa statistic (*k*)	Strength of agreement
0.00	Poor
0.00–0.20	Slight
0.21–0.40	Fair
0.41–0.60	Moderate
0.61–0.80	Substantial
0.81–1.00	Almost perfect

Table 4 Stage assignment using identified pre-defined groups and self-reports

Stage	(1) Pre-defined count	(2) Self-reported count
1: Pre-contemplation (PC)	328	308
2: Contemplation (C)	244	221
3: Preparation (P)	172	113
4: Action (A)	39	58
5: Maintenance (M)	43	126

3 Results

3.1 Evaluation of pre-defined multimodality groups mapped onto TTM stages

Table 4 displays the number of study participants assigned to the five TTM stages using the multimodality data analysed through the pre-defined approach (in

Table 5 Confusion matrix mapping the multimodality data analysed through the pre-defined approach and the self-reported TTM stages

		Self-reported TTM stages					Accuracy	Precision	Recall	F1-score	kappa coefficient
		PC	C	P	A	M					
Pre-defined approach	PC	**204**	93	23	6	2	0.7240	0.6220	0.6623	0.6415	0.4175
	C	65	**76**	52	22	29	0.6211	0.3115	0.3439	0.3269	0.0642
	P	33	39	**26**	17	57	0.7179	0.1512	0.2301	0.1825	0.0207
	A	5	10	9	**4**	11	0.8923	0.1026	0.0690	0.0825	0.0280
	M	1	3	3	9	**27**	0.8608	0.6279	0.2143	0.3195	0.2624
Overall							0.7632	0.4080	0.4080	0.4080	0.2600

column 1) and based on the self-reported data (in column 2). To evaluate the accuracy of the pre-defined approach, Table 5 presents a confusion matrix. True positives are in bold. We calculate the precision, recall, accuracy, F1-score and kappa statistic for each stage using true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN) of each stage. To estimate the overall values for each criterion we consider the sums of true TP, TN, FP, and FN.

The overall accuracy of 76% and kappa coefficient of 0.26 indicate an acceptable performance of the pre-defined approach. Notably, while the accuracies are higher than 50% (which is considered to be an acceptable level by [28]) for all the stages, kappa coefficients indicate only a “slight” performance of the pre-defined approach for all stages except the pre-contemplation and maintenance stage. Their kappa coefficient values indicate that the approach is “moderately” capable of assigning individuals to the pre-contemplation stage, and it shows “fair” capability to assign individuals to the maintenance stage. The F1 scores for these stages are 0.64 and 0.32, respectively. Given the higher accuracy, F1-score, and kappa coefficient values for these two stages, the pre-defined approach is capable of assigning individuals to these two stages more accurately than other stages.

3.2 Evaluation of the data-driven multimodality groups allocated to TTM stages

Based on the initial assessment for determining the optimum number of clusters, we use cluster groups consisting of six, eight, and ten for further analysis, while excluding three and four cluster groups due to their inability to be assigned to the five stages of TTM.

We first use the k-mean algorithm on the measures HHI, count, share, and rank to identify initial cluster groups (6 clusters, 8 clusters, and 10 clusters). Subsequently, we use the K-means algorithm considering the mean share of cycle trips variable for the derived clusters to group them into the five stages. Then we consider the mean cycling share of each group after the second clustering (6th column in Table 6). The stage assignment is based on the highest to lowest average share groups, with stages five to one being assigned accordingly (TTM stage 1-the group with the lowest mean cycling share, TTM stage 5-the group with the highest mean cycling share, and TTM stage 4,3, and 2-rest of the groups with descending order of mean cycling share). We follow the same procedure for all three cluster groups (6 clusters, 8 clusters, and 10 clusters). The accuracy of stage assignment is higher for the 6 cluster group compared with the 8 and 10 cluster groups. Therefore, we only show the results for cluster group 6 in Table 6.

Table 6 TTM stage assignment of six cluster groups identified for cycle use

Cluster number	Mean HHI	Mean count	Mean share	Mean rank	Mean share after 2nd clustering	TTM stage
1	0.3777	2.9552	0.2907	1.7015	0.2907	5
2	0.1785	4.3445	0.0564	4.4034	0.0564	2
3	0.3443	3.0000	0.0252	3.7790	0.0191	1
4	0.6903	1.7387	0.0129	2.6712		
5	0.2274	4.1714	0.1239	3.0000	0.1239	3
6	0.1677	4.4242	0.2681	1.6591	0.2681	4

Table 7 Confusion matrix for data-driven approach considering 5 stages for cycle use

		Self-reported TTM stages					Accuracy	Precision	Recall	F1-score	kappa coefficient
		PC	C	P	A	M					
Data-driven approach	PC	**224**	115	43	6	15	0.6816	0.5558	0.7273	0.6301	0.3593
	C	30	**46**	20	9	14	0.6998	0.3866	0.2081	0.2706	0.1026
	P	24	28	**27**	12	14	0.8015	0.2571	0.2389	0.2477	0.1337
	A	15	24	15	**24**	54	0.8281	0.1818	0.4138	0.2526	0.1719
	M	15	8	8	7	**29**	0.8366	0.4328	0.2302	0.3005	0.2178
Overall							0.7695	0.4237	0.4237	0.4237	0.2797

Similar to the pre-defined approach, we develop a confusion matrix to evaluate the capabilities of the data-driven approach which is described in Table 7. True positives are in bold. We estimate the precision, recall, accuracy, F1-score, and kappa coefficient values for each stage. Thereafter, we estimate the overall criteria values like for the pre-defined approach. The overall accuracy of the data-driven approach is approximately 77%, which is higher than the pre-defined approach. The overall kappa coefficient is 0.28 which is also higher than the pre-defined approach. Combined with the overall accuracy value, this indicates an acceptable performance of the data-driven approach. The F1-scores of the pre-contemplation stage (0.63) and maintenance stage (0.30) are higher compared to other stages. These values combined with kappa coefficients of 0.36 and 0.22 indicate that compared to other stages, the data-driven approach more accurately assigns individuals to pre-contemplation and maintenance stages. We made a similar observation in the pre-defined approach. However, compared to the pre-defined approach, data-driven approach accuracies, F1-scores, and kappa coefficients are slightly higher for contemplation, preparation, and action stages. This implies that the data-driven approach performs better than the pre-defined approach when identifying the individuals who belong in the middle stages.

4 Discussion

The results suggest that data on multimodality can be used to place people at different stages of the Transtheoretical model (TTM) of change in the context of a shift from car use to cycling. When considering all five stages of the TTM individually, the data-driven approach yields a comparatively higher level of accuracy in stage assignment. The accuracy of both approaches is approximately 76%. These accuracy values combined with kappa coefficients imply acceptable performances of both pre-defined and data-driven approaches.

The pre-defined approach demonstrates a relatively higher capability to assign individuals into pre-contemplation and maintenance groups. Therefore, the pre-defined approach may be preferable when targeting policy delivery specifically for individuals in the early and late stages of behaviour change which corresponds to individuals who have never considered cycling and those who actively use cycling frequently as a mode of transportation. The middle stages consist of individuals who have considered cycling, those intending to start cycling soon, and those who already use cycling as a mode of transportation. Compared to the pre-defined approach, the data-driven approach demonstrates the capability to assign individuals to the middle three stages with similar accuracies. Therefore, the data-driven approach is preferable when policy delivery is targeted at all types of individuals.

Our findings provide the policy practitioner with the opportunity to identify individuals who belong to different stages of the TTM based on objectively observable personal travel patterns and to subsequently implement the appropriate nudging intervention for each stage. This approach requires a certain level of travel data, which can be observed through the use of travel diaries. With the advancement of technology and using already available mobile phone applications, opportunities for obtaining multimodality data will increase without requiring direct engagement from the user. Therefore, the utilization of this kind of technology allows policy practitioners to obtain the necessary travel data required for establishing TTM stages of individuals based on multi-modality approaches, ultimately enabling the delivery of multiple soft interventions more effectively. For example, a mobile application could be developed to track individuals' mobility patterns and use a data-driven approach to segment users into TTM stages. Subsequently, a range of tailored travel behaviour interventions suitable for each TTM group could be implemented through the mobile application. In addition to these applications, our approach can be used to assess the ex-ante impact of different transport policies within agent-based modelling frameworks. As there's a growing interest in using behavioural stages as an outcome measure for transport policies [1, 16, 30, 32], our approach can be used to evaluate the agents' stages before and after policy implementation using travel data available and thereby assess the impact of different transport policies within the agent-based frameworks.

While the findings of our paper are promising, certain limitations need to be acknowledged. Firstly, in the data-driven approach, only four variables were used due to a lack of data availability. Future studies could expand this scope by incorporating additional variables such as entropy and trips made entirely and partially by a certain mode to more comprehensively represent multimodality and travel mode distribution. Secondly, the present study only demonstrated the TTM stage assignment in the context of a shift from driving to cycling. With additional datasets, future research could explore the efficacy of multimodality approaches for other sustainable travel modes such as public transport. Thirdly, the accuracy assessments relied on self-reported data about where respondents place themselves in the TTM. Although the two-step verification approach outperforms other methods for TTM stage assignment, it still relies on self-reported data. Thus, future studies could investigate and compare the accuracies of multimodality approaches with multiple-stage assignment tools. Lastly, our study did not account for the dynamic nature of behavioural change, particularly in terms of 'backward' or 'forward' movements on the stages of the Transtheoretical Model (TTM) since the data set we used was elicited only once, thus limiting our ability to capture changes in individuals' travel behaviour over time. This highlights an opportunity for future research to explore the longitudinal dynamics of behavioural change and its implications for sustainable mobility interventions.

5 Conclusion

This paper shows that it is possible to use multimodality data to assign people to the five stages of the Trans-Theoretical Model (TTM) of change. This assignment can be based on pre-defined categories or be data-driven relying on algorithms, and both approaches can lead to accuracies of over 75%. Accuracies are highest for the early and late stages of the TTM. We believe that assigning people to stages of the TTM using objectively measurable multimodality data will support future work that personalises soft behavioural interventions such as nudges that aim to encourage more sustainable and active travel.

Appendix A

See Table 8.

Table 8 Summary statistics of the sample

Variable	Description	Mean (S.D.)
Gender	1: if female / 0: male	0.58 (0.49)
Age	Continuous (years)	36.92 (12.85)
Education	1: if has college degree / 0: do not hold a college degree	0.70 (0.46)
Driving license	1: if holds a driving license / 0: do not have a driving license	0.93 (0.25)
Cycle ownership	1: owns a cycle / 0: do not own a cycle	0.79 (0.41)
Employment	1: work full time /0: do not work full time	0.66 (0.48)
Total trips taken	Continuous (trips)	21.85 (11.48)
Work/ school trips taken	Continuous (trips)	8.17 (6.31)
Shopping trips taken	Continuous (trips)	5.88 (4.40)
Leisure trips taken	Continuous (trips)	7.80 (5.55)

Abbreviation
TTM Trans-theoretical model

Acknowledgements
The authors would like to acknowledge and thank the Science Foundation Ireland for funding this research under the Next Generation Energy Systems (NexSys) Partnership Programme.
The authors would like to express their sincere gratitude to Associate Professor Dr Amanda Stathopoulos for sharing the survey data necessary for the present study.

Author contributions
Warnakulasooriya Umesh Ashen Lowe: Conceptualization, Visualization, Methodology, Data obtain, Formal analysis, Investigation, Writing – original draft, Leonhard Lades: Conceptualization, Visualization, Resources, Writing – review & editing, Supervision. Páraic Carroll: Conceptualization, Visualization, Resources, Writing – review & editing, Supervision.

Funding
This article has been published open access with support of the TRA2024 project funded by the European Union.

Declarations

Availability of data and materials
The primary data were obtained from a survey conducted by Biehl et al. [5]. Relevant analysis results conducted for the present study can be shared on request with the permission of the original authors (Alec Biehl, Alireza Ermagun, and Amanda Stathopoulos).

Competing interest
The authors declare that they have no known competing interests that could have appeared to influence the work reported in this paper.

Received: 11 October 2023 Accepted: 6 August 2024
Published online: 26 August 2024

References

1. Ahmed, S., Adnan, M., Janssens, D., & Wets, G. (2020). A personalized mobility-based intervention to promote pro-environmental travel behaviour. *Sustainable Cities and Society, 62*, 102397. https://doi.org/10.1016/j.scs.2020.102397
2. Aittasalo, M., Miilunpalo, S., & Suni, J. (2004). The effectiveness of physical activity counselling in a work-site setting: A randomized, controlled trial. *Patient Education and Counseling, 55*(2), 193–202. https://doi.org/10.1016/j.pec.2003.09.003
3. An, Z., Heinen, E., & Watling, D. (2023). The level and determinants of multimodal travel behaviour: Does trip purpose make a difference? *International Journal of Sustainable Transportation, 17*(2), 103–117. https://doi.org/10.1080/15568318.2021.1985195
4. Bamberg, S. (2013). Changing environmentally harmful behaviours: A stage model of self-regulated behavioural change. *Journal of Environmental Psychology, 34*, 151–159. https://doi.org/10.1016/j.jenvp.2013.01.002
5. Biehl, A., Ermagun, A., & Stathopoulos, A. (2018). Modelling determinants of walking and cycling adoption: A stage-of-change perspective. *Transportation Research Part F: Traffic Psychology and Behaviour, 58*, 452–470. https://doi.org/10.1016/j.trf.2018.06.016
6. Biehl, A., Ermagun, A., & Stathopoulos, A. (2019). Utilizing multi-stage behaviour change theory to model the process of bike share adoption. *Transport Policy, 77*, 30–45. https://doi.org/10.1016/j.tranpol.2019.02.001
7. Biondi, B., Romanowska, A., & Birr, K. (2022). Impact evaluation of a cycling promotion campaign using daily bicycle counters data: The case of Cycling May in Poland. *Transportation research part A: Policy and practice, 164*, 337–351.
8. Cardinal, B. (1997). Construct validity of stages of change for exercise behaviour. *American Journal of Health Promotion, 12*(1), 68–74.
9. Cowell, F. A. (2011). *Measuring inequality*. Oxford University Press.
10. Damgaard, C., & Weiner, J. (2000). Describing inequality in plant size or fecundity. *Ecology, 81*(4), 1139–1142.
11. Department of the Environment, Climate and Communications. (2023). *CLIMATE ACTION PLAN 2023- Changing Ireland for the better*.
12. Diana, M., & Mokhtarian, P. L. (2009). Desire to change one's multimodality and its relationship to the use of different transport means. *Transportation research part F: Traffic psychology and behaviour, 12*(2), 107–119.
13. Diana, M., & Pirra, M. (2016). A comparative assessment of synthetic indices to measure multimodality behaviours. *Transportmetrica A: Transport Science, 12*(9), 771–793.
14. Forward, S. E. (2014). Exploring people's willingness to bike using a combination of the theory of planned behaviour and the transtheoretical model. *Revue Europeenne de Psychologie Appliquee, 64*(3), 151–159. https://doi.org/10.1016/j.erap.2014.04.002
15. Friman, M., Huck, J., & Olsson, L. E. (2017). Transtheoretical model of change during travel behaviour interventions: An integrative review. *International Journal of Environmental Research and Public Health, 14*(6), 581. https://doi.org/10.3390/ijerph14060581
16. Friman, M., Maier, R., & Olsson, L. E. (2019). Applying a motivational stage-based approach in order to study a temporary free public transport intervention. *Transport Policy, 81*, 173–183. https://doi.org/10.1016/j.tranpol.2019.06.012
17. Gatersleben, B., & Appleton, K. M. (2007). Contemplating cycling to work: Attitudes and perceptions in different stages of change. *Transportation Research Part A: Policy and Practice, 41*(4), 302–312. https://doi.org/10.1016/j.tra.2006.09.002
18. Guo, B., Aveyard, P., Fielding, A., & Sutton, S. (2009). Using latent class and latent transition analysis to examine the transtheoretical model staging algorithm and sequential stage transition in adolescent smoking. *Substance Use and Misuse, 44*(14), 2028–2042. https://doi.org/10.3109/10826080902848665
19. Hausenblas, H. A., Nigg, C. R., Downs, D. S., Fleming, D. S., & Connaughton, D. P. (2002). Perceptions of exercise stages, barrier self-efficacy, and decisional balance for middle-level school students. *Journal of Early Adolescence, 22*(4), 436–454. https://doi.org/10.1177/027243102237191
20. Heinen, E. (2018). Are multimodals more likely to change their travel behaviour? A cross-sectional analysis to explore the theoretical link between multimodality and the intention to change mode choice. *Transportation Research Part F: Traffic Psychology and Behaviour, 56*, 200–214. https://doi.org/10.1016/j.trf.2018.04.010
21. Heinen, E., & Ogilvie, D. (2016). Variability in baseline travel behaviour as a predictor of changes in commuting by active travel, car and public transport: A natural experimental study. *Journal of Transport and Health, 3*(1), 77–85. https://doi.org/10.1016/j.jth.2015.11.002
22. Kormos, C., Sussman, R., & Rosenberg, B. (2021). How cities can apply behavioural science to promote public transportation use. *Behavioral Science & Policy, 7*(1), 95–115.
23. Kristal, A. S., & Whillans, A. V. (2020). What we can learn from five naturalistic field experiments that failed to shift commuter behaviour. *Nature Human Behaviour, 4*(2), 169–176. https://doi.org/10.1038/s41562-019-0795-z
24. Kroesen, M. (2014). Modelling the behavioural determinants of travel behaviour: An application of latent transition analysis. *Transportation Research Part A: Policy and Practice, 65*, 56–67. https://doi.org/10.1016/j.tra.2014.04.010
25. Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159–174.
26. McKee, R., Mutrie, N., Crawford, F., & Green, B. (2007). Promoting walking to school: Results of a quasi-experimental trial. *Journal of Epidemiology and Community Health, 61*(9), 818–823. https://doi.org/10.1136/jech.2006.048181
27. Mills, S. (2022). Personalized nudging. *Behavioural Public Policy, 6*(1), 150–159. https://doi.org/10.1017/bpp.2020.7
28. Mirzahossein, H., Bakhtiari, A., Kalantari, N., & Jin, X. (2022). Investigating mandatory and non-mandatory trip patterns based on socioeconomic characteristics and traffic analysis zone features using deep neural networks. *Computational Urban Science*. https://doi.org/10.1007/s43762-022-00063-w
29. Mizutani, S., Ekuni, D., Furuta, M., Tomofuji, T., Irie, K., Azuma, T., Kojima, A., Nagase, J., Iwasaki, Y., & Morita, M. (2012). Effects of self-efficacy on oral health behaviours and gingival health in university students aged 18- or

19-years-old. *Journal of Clinical Periodontology, 39*(9), 844–849. https://doi.org/10.1111/j.1600-051X.2012.01919.x

30. Mundorf, N., Redding, C. A., & Paiva, A. L. (2018). Sustainable transportation attitudes and health behaviour change: Evaluation of a brief stage-targeted video intervention. *International Journal of Environmental Research and Public Health, 15*(1), 150. https://doi.org/10.3390/ijerph15010150
31. Offiaeli, K., & Yaman, F. (2021). Social norms as a cost-effective measure of managing transport demand: Evidence from an experiment on the London underground. *Transportation Research Part A: Policy and Practice, 145*, 63–80. https://doi.org/10.1016/j.tra.2020.12.006
32. Olsson, L. E., Friman, M., Kawabata, Y., & Fujii, S. (2021). Integrating planned behaviour and stage-of-change into a cycling campaign. *Sustainability, 13*(18), 10116. https://doi.org/10.3390/su131810116
33. Olsson, L. E., Huck, J., & Friman, M. (2018). Intention for car use reduction: Applying a stage-based model. *International Journal of Environmental Research and Public Health, 15*(2), 216. https://doi.org/10.3390/ijerph15020216
34. Peer, E., Egelman, S., Harbach, M., Malkin, N., Mathur, A., & Frik, A. (2020). Nudge me right: Personalizing online security nudges to people's decision-making styles. *Computers in Human Behavior, 109*, 106347. https://doi.org/10.1016/j.chb.2020.106347
35. Prochaska, J. O., & Velicer, W. F. (1997). The transtheoretical model of health behaviour change. *American Journal of Health Promotion, 12*(1), 38–48.
36. Rhoades, S. A. (1993). The Herfindahl-Hirschman Index. *Federal Reserve Bulletin, 79*, 188.
37. Rosenbluth, G. (1955). Measures of concentration. In *Business concentration and price policy*, (pp. 55–99). Princeton University Press.
38. Rose, G., & Marfurt, H. (2007). Travel behaviour change impacts of a major ride-to-work day event. *Transportation Research Part A: Policy and Practice, 41*(4), 351–364. https://doi.org/10.1016/j.tra.2006.10.001
39. Semenescu, A., Gavreliuc, A., & Sârbescu, P. (2020). 30 Years of soft interventions to reduce car use—A systematic review and meta-analysis. *Transportation Research Part D: Transport and Environment*. https://doi.org/10.1016/j.trd.2020.102397
40. Skarin, F., Olsson, L. E., Friman, M., & Wästlund, E. (2019). Importance of motives, self-efficacy, social support and satisfaction with travel for behaviour change during travel intervention programs. *Transportation Research Part F: Traffic Psychology and Behaviour, 62*, 451–458. https://doi.org/10.1016/j.trf.2019.02.002
41. Spencer, L., Wharton, C., Moyle, S., & Adams, T. (2007). The transtheoretical model as applied to dietary behaviour and outcomes. *Nutrition Research Reviews, 20*(1), 46–73. https://doi.org/10.1017/S0954422407747881
42. Sunio, V., Schmöcker, J. D., & Kim, J. (2018). Understanding the stages and pathways of travel behaviour change induced by technology-based intervention among university students. *Transportation Research Part F: Traffic Psychology and Behaviour, 59*, 98–114. https://doi.org/10.1016/j.trf.2018.08.017
43. Sunstein, C. R. (2013). The Storrs lectures: Behavioral economics and paternalism. *The Yale Law Journal, 122*(7), 1867–1899.
44. Susilo, Y. O., & Axhausen, K. W. (2014). Repetitions in individual daily activity-travel–location patterns: A study using the Herfindahl-Hirschman Index. *Transportation, 41*(5), 995–1011. https://doi.org/10.1007/s11116-014-9519-4
45. Sustainable Energy Authority of Ireland. (2020). *Driving Purchases of Electric Vehicles in Ireland Behavioural insights for policy series*. www.seai.ie.
46. Tang, T., Guo, Y., Wang, H., Li, X., & Agrawal, S. (2024). Determinants of helmet use intention among E-bikers in China: An Application of the theory of planned behaviour, the health belief model, and the locus of control. *Transportation Research Record, 2678*(2), 753–769. https://doi.org/10.1177/03611981231176290
47. Thigpen, C. G., Driller, B. K., & Handy, S. L. (2015). Using a stages of change approach to explore opportunities for increasing bicycle commuting. *Transportation Research Part D: Transport and Environment, 39*, 44–55.
48. Thigpen, C., Fischer, J., Nelson, T., Therrien, S., Fuller, D., Gauvin, L., & Winters, M. (2019). Who is ready to bicycle? Categorizing and mapping bicyclists with behaviour change concepts. *Transport Policy, 82*, 11–17. https://doi.org/10.1016/j.tranpol.2019.07.011
49. Transport Scotland. (2023). *Literature review—Best practice in active travel and its associated benefits*.
50. Vij, A., Carrel, A., & Walker, J. L. (2011). Capturing modality styles using behavioural mixture models and longitudinal data. In *2nd International choice modelling conference*.
51. Wang, Z., He, S. Y., & Leung, Y. (2018). Applying mobile phone data to travel behaviour research: A literature review. *Travel Behaviour and Society, 11*, 141–155. https://doi.org/10.1016/j.tbs.2017.02.005
52. Wu, X., & Kumar, V. (2009). *The top ten algorithms in data mining* (1st ed.). CRC Press.

Publisher's Note