Journal of Transport & Health 12 (2019) 42–49

Contents lists available at ScienceDirect

Journal of Transport & Health

journal homepage: www.elsevier.com/locate/jth

Is active travel part of a healthy lifestyle? Results from a latent class analysis

Maarten Kroesen

Delft University of Technology, Faculty of Technology, Policy and Management, P.O. Box 5015, 2600 GA Delft, the Netherlands

ARTICLE INFO

Keywords:
Active travel
Walking
Cycling
Health behaviors
SNAP factors
Latent class analysis

ABSTRACT

Introduction: Behavioral health risk factors are major causes of morbidity and mortality worldwide. The four main risk factors, the so-called SNAP-factors, relate to smoking, nutrition, alcohol consumption and physical inactivity. A consistent finding in health research is that these behaviors tend to cluster together, thereby resulting in patterns of healthy lifestyles and unhealthy ones. In research to date, physical (in)activity is typically included using broad categories relating to the total amount of physical (in)activity. As such, it is unknown to what extent active travel behaviors (i.e. walking and cycling) as specific forms of physical activity are related to the health lifestyles.

Methods: In this study this knowledge gap is addressed by performing a latent class analysis based on indicators related to active travel as well as the four SNAP-factors. Data are obtained from the LISS (Longitudinal Internet Studies for the Social sciences) panel, which is based on a true probability sample of Dutch households. In total, 2050 participants are considered in the analysis.

Results: Five health lifestyles are revealed and labeled as follows: consistent healthy, active commuters, physically inactive, unhealthy eaters and consistent unhealthy. The results indicate that active travel (or lack thereof) indeed forms an integral part of the consistent healthy (and unhealthy) lifestyles. In addition, lifestyle membership is found to be significantly dependent on gender, age and level of education.

Conclusion: For most people (70%) active travel (or lack thereof) indeed forms an integral part of these consistent healthy and unhealthy lifestyles.

1. Introduction

Behavioral health risk factors are major causes of morbidity and mortality worldwide (Lim et al., 2012). The four main behavioral risk factors are smoking, poor nutrition (lack of fruit and vegetable intake), excess alcohol consumption and physical inactivity, the so-called SNAP factors. A consistent finding in health research is that these behaviors tend to co-occur or cluster together, thereby resulting in patterns of healthy lifestyles and unhealthy ones (Noble et al., 2015; McAloney et al., 2013; Meader et al., 2016). This empirical finding suggests that an approach is needed which focuses on tackling multiple risk factors simultaneously, instead of strategies that focus on changing isolated behaviors (Prochaska et al., 2008). Such an approach is further support by evidence indicating that the risk factors have synergistic effects on health, i.e. combinations of risk behaviors are more detrimental to health than their individual effects (French et al., 2008; Poortinga, 2007).

To support the development of a comprehensive strategy it is necessary to know which risk factors indeed cluster together in

E-mail address: m.kroesen@tudelft.nl.

https://doi.org/10.1016/j.jth.2018.11.006
Received 7 February 2018; Received in revised form 12 November 2018; Accepted 27 November 2018
Available online 05 December 2018
2214-1405/

particular contexts and populations. Indeed, already a fair amount of research has been devoted to this subject (Noble et al., 2015; McAloney et al., 2013; Meader et al., 2016). Based on review of 56 studies, Noble et al. (2015) found that the majority of studies (81%) reported a relatively 'healthy' cluster, which is characterized by the absence of any behavioral risk factors. In addition, half of the studies revealed the presence of a consistent unhealthy lifestyle, in which all risk factors (smoking, poor nutrition, excess alcohol consumption and high physical inactivity) were prevalent (Noble et al., 2015).

To date, physical (in)activity is typically included using broad categories relating to the amount of time spend sedentary and/or the amount of general physical activity. As such, it is unknown to what extent active travel behaviors (i.e. walking and cycling) as specific forms of physical activity are related to the different lifestyles. This study aims to address this knowledge gap and assess the extent active travel (or the lack thereof) is part of general healthy (or unhealthy) lifestyles. If active travel is indeed part of a comprehensive healthy lifestyle pattern(s), this has important policy implications, as it suggests that active travel may be increased by stimulating the more generic health lifestyles.

To reveal the health lifestyles, a latent class analysis is performed based on indicators related to active travel as well as a general measure of physical inactivity and indicators related to the three other behavioral risk factors (smoking, alcohol and nutrition). Socio-demographic variables (gender, age and education level) and the Body-Mass Index (BMI) are included as covariates to further profile the classes. The data used for the analysis are obtained from two surveys (related to health and travel behavior) conducted in the LISS (Longitudinal Internet Studies for the Social sciences) panel.[1] This panel is based on a true probability sample of Dutch households. 2050 individuals completed both surveys and are included in the analysis.

In the following, relevant empirical findings and theoretical mechanisms will be discussed. After that, the empirical study will be introduced and its results discussed. In the final section, the conclusions are summarized and several policy recommendations are formulated.

2. Empirical and theoretical background

Active travel (walking and cycling) is increasingly being recognised as a potentially effective means of increasing physical activity levels and thereby contribute to physical and mental health (Sallis et al., 2004; Frank et al., 2006; Pucher et al., 2010). Active travel can often easily be incorporated in the daily routine. In addition, there is much scope for active travel to help people meet recommended physical activity levels. In the US, for example, 27% of all trips in 2009 were shorter than 1 mile, but only 36% of those short trips were made by walking or cycling (Buehler et al., 2011).

Research related to active travel is largely driven by two questions: (1) what are the health effects of active travel? and (2) what are the causes of active travel? Multiple disciplines are involved in answering these two questions and the resulting literatures are vast. Relevant potential outcomes include increased total physical activity, reduced obesity, increased fitness and increased psychological well-being (see Oja et al., 2011, Wanner et al., 2012 and Saunders et al., 2013 for relevant reviews). Research related to the determinants of active travel has focused on the role of the built environment (e.g. residential density) and available bicycle and pedestrian infrastructure (a review of reviews is provided by Ding and Gebel (2012)). Also psychological factors (perceived environmental characteristics or attitudes and preference) have been considered, albeit to a lesser extent (Panter and Jones, 2010; Heinen et al., 2011).

Given that active travel is important health behaviour, it is relevant to know whether and to what extent it clusters with other health behaviours. As discussed above, the literature regarding the clustering of health behaviors is already quite extensive. In the health domain, many studies adopt an empirical perspective and focus on revealing the existing behavioral patterns. In this regard, consistent healthy and unhealthy clusters have been reported (Noble et al., 2015; McAloney et al., 2013; Meader et al., 2016). Moreover, it has been observed that some behaviors are more strongly correlated than others. For example, excessive alcohol use and smoking are typically found to cluster together, and the same holds for physical inactivity and poor diet (Meader et al., 2016). Other combinations are less common. For example, no association is generally found between excessive alcohol intake and physical inactivity (Noble et al., 2015).

Research has also focused on the determinants of profile membership. Here, the most consistent finding is that lower socio-economic status in terms of education level, income or type of occupation is associated with membership of more risky clusters. Regarding other socio-demographic characteristics the evidence is mixed. Males tend to have a higher probability than females of being a member of an unhealthy cluster, but in general gender is found to be a weak predictor of cluster membership. For age, the effects are also inconsistent. Although some studies report that younger age is associated with multiple risk behaviors (see e.g. Poortinga, 2007), most studies report non-significant findings (Noble et al., 2015) and some even indicate that older age is associated with riskier clusters (see e.g. Lee et al., 2012).

The theoretical mechanisms which supposedly underlie the clustering of health behaviors are still poorly understood. In the literature at least three categories of theoretical mechanisms have been identified, namely biological, psychological and sociological. Examples of biological mechanisms include notions that smoking (nicotine) may counteract the depressant effects of alcohol use or that heavy smoking may reduce lung function and thereby discourage physical activity. These biological explanations can account for the observation that some health behaviors are more strongly correlated than others.

Psychological mechanisms generally relate to the idea that health behaviors (at least partially) are the outcomes of a rational choice process, which, for example, is assumed by psychological theories such as the Health Belief Model (Janz and Becker, 1984) or

[1] www.lissdata.nl.

the Theory of Planned Behavior (Ajzen, 1991). Clustering would be expected if individuals, via such a rational process, consistently reach the same (bad/good) decisions regarding whether to adopt certain health behaviors or not.

Finally, sociological studies draw attention to broader institutional/structural explanations. An example is the health lifestyle theory developed by Cockerham (2005). This theory postulates that particular health lifestyles originate from the interplay of life choices (agency) and life chances (structure). The life chances are, amongst others, determined by class circumstances and living conditions, which may operate as either constraints or enablers of certain health lifestyles, resulting in consistent healthy or unhealthy behavioral patterns. Hence, sociological theories (like the health lifestyle theory) emphasize the fact that some groups in society have fewer life chances than others (due to a lack of resources, time, or access to healthy food and exercise opportunities) and can therefore be expected to engage in multiple health risk behaviors.

While empirical research so far has shed little light as to which theory (or combination of theories) actually explains the clustering of health behaviors, it concluded that various mechanisms can be identified that may be responsible. Yet, in future work, it would be interesting to try and uncover which mechanisms are actually most relevant. This issue will be returned to in the discussion.

Both psychological and sociological theories ascribe an important role to the level of education as a determinant of engaging in multiple health behaviors. Within psychological models, it may be expected that higher educated people have more knowledge of health behaviors (and can also more easily acquire new knowledge), which arguably will result in better (healthier) decisions. In sociological models, the level of education is an important indicator of class (in postmodern societies) which, as discussed above, may enable or constrain certain health behaviors. For example, higher educated people may engage in healthy lifestyle, because it is perceived as the class norm, thereby also allowing them to distinguish themselves from other ('lower') classes.

This study contributes to the health literature by empirically assessing to what extent active travel (or lack thereof), as a specific form of physical (in)activity, is part of (un)healthy lifestyles. Typically, broad measures of physical activity are considered, relating to the total amount of sedentary time or general physical activity. As such, it is as of yet unknown to what extent forms of active travel are related to health lifestyles.

This study also aims to contribute to the transportation literature. In transportation research, active travel is often (implicitly) conceptualized as an 'environmentally friendly' behavior, as opposed to a 'health-enhancing' behavior. As such, research typically considers psychological constructs related to environmental attitudes or beliefs to explain active travel (see e.g. Bamberg and Möser, 2007). While correlations between such attitudes and active travel are typically found, it might be that (some) individuals also engage in active travel as a means to stay fit and healthy.

One way to test this idea is to measure both of these psychological motivations and assess their effects on active travel. Empirically, however, as shown by the research of Heinen et al. (2011), motivations related to environmental and health benefits are highly intercorrelated (both motivations were actually found to load on the same factor). Hence, it has proven difficult to assess the (unique) contribution of health motivations in explaining variation in active travel. In addition, the motivations may also be adopted post-hoc, as a way to justify the behavior (Kroesen et al., 2017). Hence, the direction of causation always remains uncertain.

By assessing the extent to which active travel is part of healthy lifestyles this problem can be addressed to some extent. More specifically, should active travel (indeed) occur mainly within comprehensive healthy lifestyles, this would suggest that health motivations indeed play a relevant role (in addition to the possible role of environmental motivations). While such evidence does not prove that health motivations do play a role, any lack of clustering of active travel and other health behaviors would quite definitively prove that health motivations do not play a role. So the present study should be regarded as another relevant piece of the puzzle with regard to the role of health motivations in determining level of active travel.

3. Method

3.1. Data and measures

The data used for the analysis are drawn the LISS (Longitudinal Internet Studies for the Social sciences) panel, which is based on a true probability sample of Dutch households.[2] From this panel data from two surveys are combined, one survey on travel behavior conducted in July 2013[3] and one on health conducted in November 2013.[4] For the travel behavior survey 2980 panel members were invited and 2370 responded (response rate 79.5%) and for the health survey 6217 were invited and 5379 responded (response rate 79.5%). In total, 2050 individuals completed both of these surveys and are included in the analysis.

Since (additional) selection bias may have been introduced by considering only people who participated in both surveys (although both individually have quite high response rates), several representativeness tests were conducted. Table 1 presents the sample distributions of three sociodemographic variables (gender, age and level of education) and Body-Mass Index (BMI) and the respective population distributions retrieved from Statistics Netherlands. The results indicate that the sample is representative for the population of Dutch adults with respect to gender and BMI (no significant differences). However, the mean age in the sample is 3.9 years higher than in the population, and also the level of education is (on average) higher in the sample compared to the population. The implications of these results with respect to the findings will be reflected upon in the concluding section.

Active travel was operationalized using three indicators: the distance travelled by bicycle in a regular week (measured on a 5-

[2] Details on the panel can be found at www.lissdata.nl.

[3] See https://www.dataarchive.lissdata.nl/study_units/view/584.

[4] See https://www.dataarchive.lissdata.nl/study_units/view/509.

Table 1
Sample and population distributions of sociodemographic variables and BMI.

		Sample	Population[a]	Test of significance
Gender (%)	Male	47	49	χ^2 = 2.3, *df* = 1, p-value = 0.127
	Female	53	51	
Age	Mean	51.6	47.7	t = 10.5, *df* = 2049, p-value = 0.000
Level of education (%)	Low	32	33	χ^2 = 3.5, *df* = 2, p-value = 0.000
	Intermediate	36	40	
	High	33	27	
BMI (kg/m^2) (%)	Underweight (< 18.5)	2	2	χ^2 = 6.5, *df* = 3, p-value = 0.090
	Normal weight (18.5–25)	48	50	
	Overweight (25–30)	37	36	
	Obese (> 30)	13	12	

[a] Data retrieved from Statistics Netherlands (http://statline.cbs.nl/Statweb/).

point ordinal scale), the number of days that the respondent spent at least 10 min walking in the past week and a dummy indicating whether the respondent is an active commuter or not (walking or cycling to work or school).

The SNAP-factors were measured as follows. Smoking was operationalized using a simple indicator whether the respondent currently uses tobacco (smoking cigarettes, cigars or pipes) or not. Excessive alcohol consumption was operationalized with the following question: how often did you have a drink containing alcohol over the last 12 months? Respondents could indicate their answer on an 8-point ordinal scale ranging from (1) not at all over the last 12 months to (8) almost every day. Only those falling in the highest category (8) were considered as excessive drinkers. Hence, the original scale was recoded into a binary variable indicating excessive consumption or not. The nutrition factor was operationalized using two questions relating to fruit and vegetable consumption. For these questions 6-point ordinal scales were used ranging from (1) never to (6) every day. Here, the last two categories, (5) 5–6 times per week and (6) every day, were considered as indicative of a good diet. Hence, these two variables were also recoded into binary variables. Finally, in addition to the indicators related to active travel, a broad measure of physical inactivity was included, namely the number of hours spend on sedentary activities on a regular day. For this question a 4-point ordinal scale was used ranging from (1) 0–3 h per day to (4) 10 or more hours per day. In total, eight indicators were used in the analysis.

Note that the used indicators for active travel capture active travel for both transportation and leisure purposes. Given that general physical activity is operationalized as the time spend sedentary (as opposed to some form of physical activity), there is no overlap with this indicator.

Unfortunately, the used indicators were not operationalized in such a way that general recommended thresholds for healthy (unhealthy) behaviors could be adopted. For example, WHO's recommendation regarding vegetable/fruit consumption consists of eating 400 g of fruits and vegetables per day. In a similar fashion, regarding physical activity, WHO recommends that adults aged 18–64 should do at least 150 min of moderate-intensity aerobic physical activity throughout the week. Since the scales based upon which these thresholds are defined do not align with the answering scales used in the survey, the behavioral patterns (resulting from the analysis) cannot be interpreted in an absolute sense (i.e. based on the accepted guidelines), as being either 'healthy' or 'unhealthy'. Instead, the resulting patterns should be interpret in a relatively sense, as being relatively more or less health than other patterns and/or the sample average.

Next to the indicators, the socio-demographic characteristics and BMI were included as covariates in the model. To assess the significance of these covariates the 3-step procedure was applied (Vermunt, 2010). An advantage of this procedure over the 1-step approach (where covariates are directed included to predict class membership) is that the covariates will not interfere with the measurement part of the model, i.e. classification is solely based on the indicators (and not the covariates). The 3-step procedure basically consists of the following steps: (1) estimation of the model based on the indicators only, (2) probabilistic assignment of subjects to latent classes (the posterior membership probabilities) and (3) estimation of the effects of the covariates on latent class membership, corrected for the classification error to prevent bias. The procedure allows the researcher to establish the effects (and significance) of the covariates (corrected for measurement errors), while not letting the covariates interfere with the classification based on the indicators (Vermunt, 2010).

3.2. Model estimation

Latent class modelling has several advantages over traditional clustering techniques, such as K-means cluster analysis (Magidson and Vermunt, 2002). One particular advantage is that nominal and ordinal indicators can be used (in addition to continuous outcomes), which were also present in the current application. Within the model specification all indicators were specified as either nominal or ordinal. Latent Gold 5.1 was to estimate the latent class models (Vermunt and Magidson, 2013).

The goal of the latent class analysis is to find the most parsimonious model, i.e. with the smallest number of latent classes, which adequately describes the associations between the indicators. To identify the optimal model, subsequent models were estimated with 1 through 8 latent classes. Table 2 presents the fit of these models in terms of the Bayesian information criterion (BIC), a statistic which weighs model fit and model parsimony, and the sum of the bivariate residuals (BVRs), indicating the total amount of association remaining between the indicators after accounting for the latent class variable (Vermunt and Magidson, 2013).

Based on the BIC statistic (being lowest in the 3-class model) the 3-class model should be considered optimal. However, in this

Table 2
Model fit of the latent class models.

No. of classes	Npar	LL	BIC(LL)	Sum of BVRs
1	19	−15562.6	31,270.0	492.4
2	28	−15447.4	31,108.3	261.9
3	37	−15368.7	31,019.6	85.7
4	46	−15343.3	31,037.3	63.8
5	55	−15318.8	31,057.0	26.1
6	64	−15302.3	31,092.7	25.5
7	73	−15289.1	31,134.8	13.8
8	82	−15273.8	31,172.9	7.6

Npar = number of model parameters.
LL = final log-likelihood.
BIC(LL) = Bayesian information criterion (based on log-likelihood).
Total BVR = sum of the bivariate residuals.

solution, significant bivariate residuals (> 3.84) remained between the indicators (note that since the bivariate residuals are chi-squared distributed with one degree of freedom, a value of 3.84 corresponds to the critical chi-square value at the 5% level of significance) (Vermunt and Magidson, 2013). These bivariate residuals were not reduced up until the 5-class solution. Since the 5-class solution also provided additional relevant substantive insights over the 3-class model, the decision was made to consider this solution as optimal. In the next section this solution will therefore be interpreted substantively.

4. Results

Table 3 presents the class sizes and the profiles of the five classes. To aid the interpretation the final column presents the sample distributions. Overall, the five classes represented well-interpretable patterns. Note that, although the estimation of the measurement model (model with only indicators) and structural model (model with covariates) occurs consecutively (in line with the 3-step

Table 3
Profiles of the 5 latent classes and the sample distributions.

		1	2	3	4	5	Sample
Cluster Size (%) N = 2050		41	19	18	11	10	
Indicators							
Distance (in kilometer) travelled by bicycle in a regular week (%)	0	6	3	37	20	34	16
	1–10	24	14	46	43	46	30
	11–20	19	15	11	18	12	16
	21–40	23	25	4	11	5	17
	> 40	28	44	2	8	2	21
No. of days with more than 10 min walking in past week	Mean	3.4	3.9	1.9	2.4	1.9	3.0
Active commuter (%)	No	75	28	100	72	92	72
	Yes	25	72	0	28	8	28
No. of hours sedentary on a regular day (%)	0–3	21	11	9	23	18	17
	4–6	43	34	31	43	41	39
	7–9	20	25	26	20	22	22
	10 or more	16	29	34	14	19	22
Currently smokes tobacco (%)	No	95	83	95	90	4	83
	Yes	5	17	5	10	96	17
Excessive drinking (drinking almost every day in past year) (%)	No	84	90	83	97	70	85
	Yes	16	10	17	3	30	15
Frequency of eating vegetables (5 times per week or more) (%)	No	23	34	26	95	43	36
	Yes	77	66	74	5	57	64
Frequency of eating fruit (5 times per week or more) (%)	No	15	57	36	86	62	40
	Yes	85	43	64	14	38	60
Covariates							
Gender (%) (Wald = 28.4, p = 0.00)	Male	40	50	55	56	48	47
	Female	60	50	45	44	52	53
Age (Wald = 121.0, p = 0.00)	Mean	61.9	39.5	49.3	42.5	50.6	51.6
Level of education (%) (Wald = 56.9, p = 0.00)	Low	38	20	24	40	38	32
	Intermediate	29	44	32	44	40	36
	High	33	36	43	17	22	33
BMI (kg/m^2) (%) (Wald = 21.5, p = 0.04)	Underweight (< 18.5)	1	5	0	1	2	2
	Normal weight (18.5–25)	45	60	41	45	51	48
	Overweight (25–30)	43	25	37	40	34	37
	Obese (> 30)	11	10	22	14	14	13

Note: some column values may not add up to 100% due to rounding.

procedure), the results of the third step are already included here. Hence, the distributions and tests of significance of the covariates are included in the profile output (at the bottom).

The first two classes represent relatively healthy lifestyles. Subjects belonging to the first class (41% of the sample) cycle and walk above the sample average. A substantial portion of the subjects is also an active commuter (25%). Compared to the other classes, the time spend sedentary (on average per day) is among the lowest in this class. Relatively few engage in smoking (5%) and drinking (16%) and the frequencies of eating vegetables and fruit (5 times or more per week) are high (77% and 85% respectively). Overall, the first class represents a relatively *consistent healthy* lifestyle.

While the second class (19% of the sample) also represents a relatively healthy lifestyle, there are several distinct differences with the first class. Firstly, compared to the first class, levels of cycling and walking are higher in this class. In addition, 75% of the subjects are *active commuters*. At the same time, the amount of sedentary time is also higher. Thirdly, compared to the first class, smoking occurs relatively more frequently (17%), while drinking occurs less frequently (10%). Overall, however, these levels are still at the low end of the spectrum. A similar pattern occurs with respect to vegetable and fruit intake, which is still quite high, but again slightly lower than in the first class.

The third class (18% of the sample) scores especially poor on the indicators of *physical inactivity*. Compared to the other classes, the levels of cycling and walking are lowest, while the time spent sedentary is the highest. The levels of smoking, drinking and vegetable/fruit intake are, however, comparable to the first (healthy) class. Hence, only in terms of physical inactivity does this pattern represent an unhealthy lifestyle.

The fourth class (11% of the sample) has low levels of cycling and walking, but still a substantial portion of active commuters (28%). In addition, the amount of sedentary time is relatively low and smoking/drinking also occurs relatively infrequently. The distinct feature of this class is the low level of vegetable and fruit intake; 95% and 86% respectively does not meet the threshold of eating vegetables/fruits 5 times or more per week. Hence, people with this pattern can be described as *unhealthy eaters*.

Finally, the last class (10% of the sample) represents a relatively *consistent unhealthy* lifestyle. Subjects in this class engage little in active travel. In addition, smoking and drinking levels are high (96% and 30% respectively), while the levels of fruit and vegetable intake are low.

All four covariates were found to be significant ($p < 0.05$). The results indicate that females are relatively more likely to belong to the first ('consistent healthy') lifestyle, while males are more likely to belong to the fourth ('low level of vegetable and fruit intake') class. Subjects in the first class ('consistent healthy') are on average relatively old (mean = 61.9), while subjects of the second ('active commuters') and fourth ('low level of vegetable and fruit intake') class are relatively young (mean ~ 40). Probably, the lack of active commuting in the first 'consistent healthy' class is (partly) due to the high average age in the class, resulting in the fact that relatively many in the class do not have to commute for work. Overall, the education level is higher in the healthier classes (1 and 2) compared to the unhealthier ones (4 and 5). Interesting, however, the education level is highest in the third class ('high physical inactivity'), which is probably due to the fact that people in this class are more likely to have an (sedentary) office job. In line with the level of physical activity, the BMI is highest in the third class (22% obese), followed by the relatively unhealthy classes (class 4 and 5, ~ 14% obese) and the healthier ones (class 1 and 2, ~ 10% obese).

A more intuitive way to interpret the effects of the covariates, which is also more in line with the underlying conceptualization that the covariates influence class membership, is to calculate and assess the predicted class membership probabilities for various levels of the covariates (while holding the other covariates at their mean value). This is done for age and level of education, which

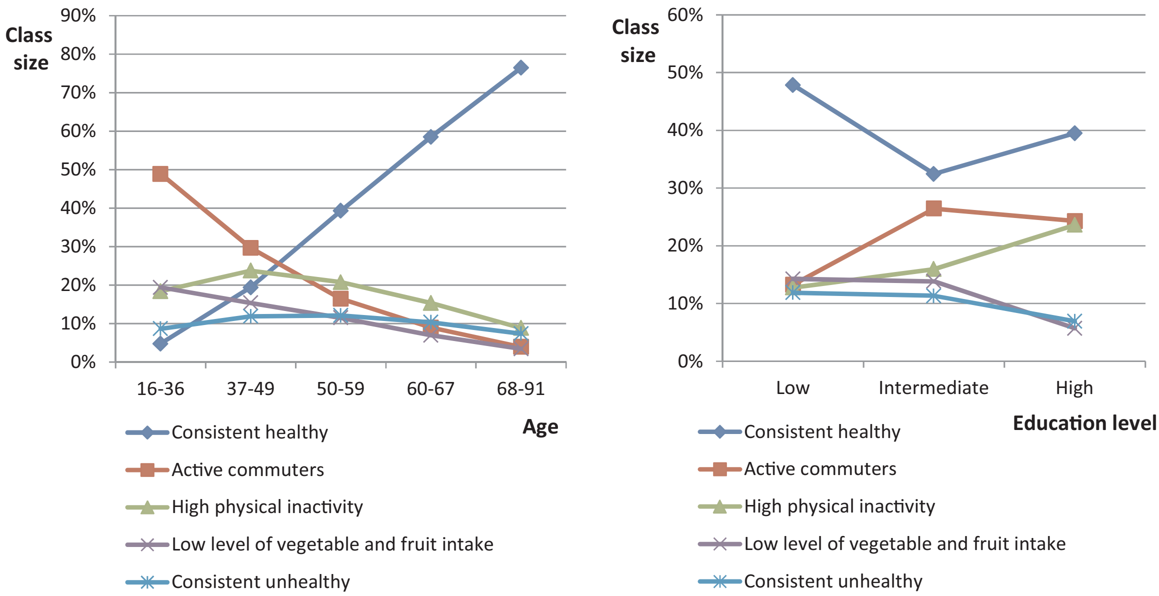

Fig. 1. Predicted class membership probabilities conditional on age (left) and education level (right).

have the strongest effects on class membership. Fig. 1 plots class membership as a function of these variables.

With respect to age it becomes clear that, over the life course, membership of the consistent healthy lifestyle (class 1) steadily increases. This growth occurs mainly at the expense of the active commuter lifestyle (class 2), but also of the other three lifestyles. Hence, with increasing age people mainly transition from the active commuter lifestyle to the consistent healthy lifestyle, thereby reducing their active travel but generally improving their health behaviors in terms of (not) smoking and their diet. This reduction in active travel is plausible given the increasing difficulty of traveling actively at older age.

The effects of level of education are somewhat inconsistent with previous research. Typically, higher education is found to be associated with healthier lifestyles (Noble et al., 2015). Indeed, the findings here also show that the probability of being a member of the consistent unhealthy lifestyle is lowest (7%) when the level of education is highest. Yet, as discussed above, membership of the relatively unhealthy 'physical inactive' profile also increases consistently with the level of education. Another interesting finding is that the level of education strongly influences the class membership distributions of the first two classes, i.e. the consistent healthy and the active commuter lifestyle. Especially in moving from the low to the intermediate level of education a large shift in these classes can be observed, whereby the active commuter pattern strongly increases at the expense of the consistent healthy lifestyle. Hence, in the Dutch context, active commuting is strongly linked to the level of education.

5. Discussion

Taken together the results of the research are in line with previous findings, although some particular unexpected results are revealed as well. Firstly, in line with previous studies, at both ends of the spectrum evidence of clustering is found. Around 60% of the sample (class 1 and 2) has a consistent healthy lifestyle, while 10% has a consistent unhealthy lifestyle (class 5). In addition, in line with the main objective of the present study, it is shown that for most people (70%) active travel (or lack thereof) indeed forms an integral part of these consistent healthy and unhealthy lifestyles, i.e. with high engagement in the overall healthy lifestyles (class 1 and 2) and low engagement in the overall unhealthy lifestyle (class 5). Yet, two classes (3 and 4) are revealed for which this does not hold. Especially the third class stands out in this regard, as it represents an overall healthy lifestyle with the exception of physical activity (including active travel).

The effects of the covariates are also in line with findings of previous studies, i.e. on a whole, the probability of being a member of one of the healthier classes increases with being female, level of education and age. Still, here as well, some unexpected findings are revealed. For example, the finding that the level of education is highest (on average) in the relatively unhealthy 'physical inactivity' profile (class 3). Another peculiar finding is that age does not strongly affect the probability of being a member of this class and an also the effect of the level of education is relatively weak. This somewhat contrasts previous research which reported stronger relationships between health lifestyles and (especially) the level of education. It may be speculated that the level of education in general is quite high in the Netherlands and in this sample in particular (as the representativeness analysis has shown), thereby reducing variation in the variable and weakening the effect on the health lifestyles.

Finally, some limitations of this study and related avenues for future research can be identified. The most important limitation is that the health risk behaviors are based on self-report, which, due to social desirability bias, tend to underestimate the prevalence of health risk behaviors and overestimate healthy behaviors (Newell et al., 1999). Yet, in many cases they remain the only feasible method of collection (Noble et al., 2015). A related issue, is that the answering scales did not match those used in setting general guidelines as to what constitutes healthy behavior (or not), a point which was discussed in section 3. Given that this is the case, it is difficult to interpret the patterns in an absolute sense, i.e. as being either healthy or not. Ideally, future research should be based on more accurate and objective measurements of active travel and the SNAP factors. For example, including measurements on actual fruit and vegetable intake or on the actual time spend on active travel. Considering the measurement of active travel in particular, in the transportation studies this is often measured using travel diaries. Probably, this approach will yield more reliable estimates of active travel compared to general questions about (weekly) travel behavior by various modes. It would be interesting to use data from such studies in future health research.

A second limitation relates to the sample representativeness and the generalizability of the findings. The representativeness analysis has shown that older and higher educated people are overrepresented in the sample compared to the population. Since age and education level positively influence the probability of being a member of the 'consistent healthy' class, this class is probably overrepresented in the sample compared to the population. At the same time, it is reassuring to see that no bias exists with respect to BMI, suggesting that no health-related selection mechanisms were at work (i.e. those with poor health being more inclined to deny participation).

A third limitation relates to the cross-sectional nature of the data, making it impossible to draw causal inferences and/or assess intra-individual change over time. This limitation may be addressed in future work as data are available from multiple waves (years) in the LISS panel. Specifically, a latent transition model (Collins and Lanza, 2013) may be estimated to reveal how people transition between the different latent classes over time and assess which factors and/or events trigger transitions to (un)healthy lifestyles.

Finally, similar to most studies in the health literature which adopt an empirical/descriptive approach, i.e. focused on revealing which clusters exist in the population, this study does not provide an answer to the question why these clusters exist. While qualitative research may shed light on this question, a quantitative approach would be to additionally include various (stated) motivations for engaging in the respective health behaviors in the latent class model. It may then be assessed what are the primary motivations for (not) engaging in multiple health behaviors. Yet, at this point, one (again) runs into the trouble that motivations may be adopted post-hoc instead of driving the behaviors.

Acknowledgement

The travel behavior survey used in the present study was developed jointly with Professor Bert van Wee and Professor Susan Handy. I am grateful for their support in developing this survey. The LISS panel data were collected by CentERdata (Tilburg University, The Netherlands) through its MESS project funded by the Netherlands Organization for Scientific Research.

Conflict of interest

None.

References

Ajzen, I., 1991. The theory of planned behavior. Organ. Behav. Hum. Decis. Process. 50 (2), 179–211.
Bamberg, S., Möser, G., 2007. Twenty years after Hines, Hungerford, and Tomera: a new meta-analysis of psycho-social determinants of pro-environmental behaviour. J. Environ. Psychol. 27 (1), 14–25.
Buehler, R., Pucher, J., Merom, D., Bauman, A., 2011. Active travel in Germany and the US: contributions of daily walking and cycling to physical activity. Am. J. Prev. Med. 41 (3), 241–250.
Cockerham, W.C., 2005. Health lifestyle theory and the convergence of agency and structure. J. Health Social Behav. 46 (1), 51–67.
Collins, L.M., Lanza, S.T., 2013. Latent Class and Latent Transition Analysis: With Applications in the Social, Behavioral, and Health Sciences John Wiley & Sons, Hoboken, New Jersey.
Ding, D., Gebel, K., 2012. Built environment, physical activity, and obesity: what have we learned from reviewing the literature? Health Place 18 (1), 100–105.
Frank, L.D., Sallis, J.F., Conway, T.L., Chapman, J.E., Saelens, B.E., Bachman, W., 2006. Many pathways from land use to health: associations between neighborhood walkability and active transportation, body mass index, and air quality. J. Am. Plan. Assoc. 72 (1), 75–87.
French, S., Rosenberg, M., Knuiman, M., 2008. The clustering of health risk behaviours in a Western Australian adult population. Health Promot. J. Aust. 19 (3), 203–209.
Heinen, E., Maat, K., Van Wee, B., 2011. The role of attitudes toward characteristics of bicycle commuting on the choice to cycle to work over various distances. Transp. Res. Part D: Transp. Environ. 16 (2), 102–109.
Janz, N.K., Becker, M.H., 1984. The health belief model: a decade later. Health Educ. Q. 11 (1), 1–47.
Kroesen, M., Handy, S., Chorus, C., 2017. Do attitudes cause behavior or vice versa? An alternative conceptualization of the attitude-behavior relationship in travel behavior modeling. Transp. Res. Part A: Policy Pract. 101, 190–202.
Lee, Y., Back, J.H., Kim, J., Byeon, H., Kim, S., Ryu, M., 2012. Clustering of multiple healthy lifestyles among older Korean adults living in the community. Geriatr. Gerontol. Int. 12 (3), 515–523.
Lim, S.S., Vos, T., Flaxman, A.D., Danaei, G., Shibuya, K., Adair-Rohani, H., Aryee, M., 2012. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 380 (9859), 2224–2260.
Magidson, J., Vermunt, J., 2002. Latent class models for clustering: a comparison with K-means. Can. J. Mark. Res. 20 (1), 36–43.
McAloney, K., Graham, H., Law, C., Platt, L., 2013. A scoping review of statistical approaches to the analysis of multiple health-related behaviours. Prev. Med. 56 (6), 365–371.
Meader, N., King, K., Moe-Byrne, T., Wright, K., Graham, H., Petticrew, M., Sowden, A.J., 2016. A systematic review on the clustering and co-occurrence of multiple risk behaviours. BMC Public Health 16 (1), 657.
Newell, S.A., Girgis, A., Sanson-Fisher, R.W., Savolainen, N.J., 1999. The accuracy of self-reported health behaviors and risk factors relating to cancer and cardiovascular disease in the general population: a critical review. Am. J. Prev. Med. 17 (3), 211–229.
Noble, N., Paul, C., Turon, H., Oldmeadow, C., 2015. Which modifiable health risk behaviours are related? A systematic review of the clustering of Smoking, Nutrition, Alcohol and Physical activity ('SNAP') health risk factors. Prev. Med. 81, 16–41.
Oja, P., Titze, S., Bauman, A., de Geus, B., Krenn, P., Reger-Nash, B., Kohlberger, T., 2011. Health benefits of cycling: a systematic review. Scand. J. Med. Sci. Sports 21 (4), 496–509.
Panter, J.R., Jones, A., 2010. Attitudes and the environment as determinants of active travel in adults: what do and don't we know? J. Phys. Act. Health 7 (4), 551–561.
Poortinga, W., 2007. The prevalence and clustering of four major lifestyle risk factors in an English adult population. Prev. Med. 44 (2), 124–128.
Prochaska, J.J., Spring, B., Nigg, C.R., 2008. Multiple health behavior change research: an introduction and overview. Prev. Med. 46 (3), 181–188.
Pucher, J., Buehler, R., Bassett, D.R., Dannenberg, A.L., 2010. Walking and cycling to health: a comparative analysis of city, state, and international data. Am. J. Public Health 100 (10), 1986–1992.
Sallis, J.F., Frank, L.D., Saelens, B.E., Kraft, M.K., 2004. Active transportation and physical activity: opportunities for collaboration on transportation and public health research. Transp. Res. Part A: Policy Pract. 38 (4), 249–268.
Saunders, L.E., Green, J.M., Petticrew, M.P., Steinbach, R., Roberts, H., 2013. What are the health benefits of active travel? A systematic review of trials and cohort studies. PLoS One 8 (8), e69912.
Vermunt, J.K., 2010. Latent class modeling with covariates: two improved three-step approaches. Polit. Anal. 18 (4), 450–469.
Vermunt, J.K., Magidson, J., 2013. Technical Guide for Latent GOLD 5.0: Basic, Advanced, and Syntax. Statistical Innovations Inc., Belmont, MA.
Wanner, M., Götschi, T., Martin-Diener, E., Kahlmeier, S., Martin, B.W., 2012. Active transport, physical activity, and body weight in adults: a systematic review. Am. J. Prev. Med. 42 (5), 493–502.