Taylor & Francis
Taylor & Francis Group

Children's school commuting in the Netherlands: Does it matter how urban form is incorporated in mode choice models?

Marco Helbich

Department of Human Geography and Spatial Planning, Utrecht University, Utrecht, the Netherlands

ABSTRACT

Walking and cycling to school is environmentally sustainable and increases children's physical activity. Although it is known that the built environment influences children's travel behavior, there is scant knowledge of how urban form along children's commuting routes affects walking and cycling separately, or of how to incorporate urban form exposures in transport mode choice models. This research investigated (a) the associations between children's transport mode choice and urban form correlates *en route*, and (b) the consequences of different urban form operationalizations (i.e., individual variables, composite measure, and principal components [PCs]). Global Positioning System devices were used to track 623 trips to and from school made by Dutch children aged 6–11 years. Urban form exposures were derived with geographic information systems, and their relationships with mode choice were tested with mixed multinomial logit models in a cross-sectional research design. Differences between the number of associated urban form variables, their magnitudes, and their significance levels were found for both walking and cycling, independent of the operationalization. Urban form was most influential for walking, whereas distance-related effects were absorbed when modeled as PCs. The highest model fit was achieved through PCs; the composite measure resulted in the lowest fit. To maximize the effectiveness of planning and health interventions, walking and cycling must be targeted separately. Policymakers should avoid "one-fits-all policies," which are deemed to be inefficient. Because urban form affects model interpretation and fit, careful attention should be paid to how urban form is modeled, and sensitivity analyses should be performed.

1. Introduction

Although children's physical activity has been declining significantly for many years (Sallis et al., 2015), physical activity is a crucial means to reduce the risk of adverse health conditions in children (Bauman et al., 2012). In the Netherlands, only 18% of children perform the recommended 60 min of physical activity a day (Hildebrandt et al., 2013). Sufficient physical activity is particularly relevant for young age cohorts, since healthy behaviors during childhood translate into healthier lifestyles in adulthood (Janssen & Leblanc, 2010).

School commuting provides a vital opportunity to increase children's energy expenditure (Faulkner, Buliung, Flora, & Fusco, 2009; Bauman et al., 2012; Larouche, Saunders, Faulkner, Colley, & Tremblay, 2014; Duncan et al., 2015). Besides individual health benefits, walking and cycling as a utilitarian source of transport induce environmental benefits (e.g., a carbon-neutral transport mode, reduction of traffic congestions during peak hours around schools) (Fishman, Böcker, & Helbich, 2015; Nieuwenhuijsen, Khreis, Verlinghieri, & Rojas-Rueda, 2016; van Wee & Ettema, 2016). Despite these associated co-benefits, the number of children who engage in active school traveling has reduced (Ferrari & Green, 2013),

while chauffeuring is on the rise in the Netherlands and other countries (Fyhri & Hjorthol, 2009; Carver, Timperio, & Crawford, 2013).

Commuting mode choice is a complex decision-making process involving both parents and children (Mitra, 2013; Zhou, 2016). In addition to children's individual characteristics (e.g., age, gender) (Steinbach, Green, & Edwards, 2012; Kaplan, Nielsen, & Prato, 2016), trip-specific issues also matter (Pont, Ziviani, Wadley, Bennett, & Abbott, 2009). In particular, the commuting distance is inversely associated with mode choice: Children who live closer to school have a higher likelihood of walking and cycling there (Ewing, Schroeer, & Greene, 2004; Yeung, Wearing, & Hills, 2008; Larson et al., 2009). Literature also indicates that stranger danger, traffic safety, and traffic speed influence mode choice (Pont et al., 2009; Carver et al., 2013). For example, not having to cross major roads and the availability of sidewalk and bike infrastructure are supportive for safe walking and cycling (Kerr et al., 2006; Helbich et al., 2016).

Meta-analysis (Ewing & Cervero, 2010) and reviews on children's travel behavior indicate associations with urban form (Pont et al., 2009; Van Loon & Frank, 2015). Because urban form

CONTACT Marco Helbich ✉ m.helbich@uu.nl ✎ Department of Human Geography and Spatial Planning, Utrecht University, Heidelberglaan 2, 3584 CS Utrecht, the Netherlands.

encompasses a myriad of components (Cervero & Kockelman, 1997; Brownson, Hoehner, Day, Forsyth, & Sallis, 2009), studies (Panter, Jones, Van Sluijs, & Griffin, 2010; Wong, Faulkner, & Buliung, 2011) frequently operationalize it through density, diversity, and design indicators (Ewing & Cervero, 2010). In theory, density reduces distances between places, which stimulates both walking and cycling (Van Loon & Frank, 2015). Land-use diversity is also grounded on a similar distance-based notion. Saelens and Handy (2008) argue that a pronounced land-use mix brings destinations closer together, and is associated with less chauffeuring. However, findings concerning land-use and building diversity are conflicting for children's transport behavior (Larson et al., 2009). Urban design is represented by means of the street layout (Giles-Corti et al., 2011; Koohsari et al., 2016). Here, the assumption is that dense street patterns characterized by high junction densities or a large number of well-connected streets support walking and cycling by increasing route choice and accessibility (Schlossberg, Greene, Phillips, Johnson, & Parker, 2006; Helbich et al., 2016). Although well-connected streets are associated with reduced automobile speeds, children need to cross streets more often, increasing the risk of traffic accidents (Sirard & Slater, 2008). Cul-de-sacs, in contrast, weaken network connectivity, impeding walkability but not necessarily cyclability (Ferrari & Green, 2013). Because these measures are limited to the horizontal dimension of urban form, Helbich et al. (2016) extended urban morphology measures to the third dimension, where the urban topography is represented through laser scanning data (Helbich, Jochem, Mücke, & Höfle, 2013). As argued, building height differences might stimulate active commuting. However, there is still controversy over the associations with and the extent of urban form on children's mode choice (Sirard & Slater, 2008; Pont et al., 2009; Mitra, 2013; Van Loon & Frank, 2015).

There are three reasons why the relationships are not universally found. The first reason is methodological, that is, the way urban form exposures are modeled. While exposures are primarily restricted to areas either around the home or the school location, or both (Larsen et al., 2009; Panter et al., 2010), *en route* exposures appear increasingly as significant determinants of travel behavior (Badland, Schofield, & Garrett, 2008; Duncan, Badland, & Mummery, 2009; Böcker, Dijst, Farber, & Helbich, 2015; Helbich et al., 2016). Straight-line distances (Steinbach et al., 2012; Owen et al., 2012) or shortest path distances (Müller, Tscharaktschiew, & Haase, 2008) between home and school are utilized to mimic the actual path. Both approximations are questionable, however, as the Euclidean distance tends to underestimate actual distances (Stigell & Schantz, 2011) and the shortest routes are considered insufficiently accurate (Krenn, Oja, & Titze, 2014; Harrison, Burgoine, Corder, Van Sluijs, & Jones, 2014). To circumvent this shortcoming, the Global Positioning System (GPS) (Kerr, Duncan, & Schipperjin, 2011), in combination with geographic information systems (GIS), is increasingly used to model route-specific exposures (Park & Kwan, 2017).

Second, apart from a few exceptions (Schlossberg et al., 2006; Steinbach et al., 2012), different modes are not analyzed separately. It is common practice to combine walking and cycling into a single category called active modes (Bringolf-Isler et al., 2008; Larson et al., 2009; Broberg, Salminen, & Kyttä, 2013). However, as

highlighted by several authors (Giles-Corti, Timperio, Bull, & Pikora, 2005; Wong et al., 2011; Lu et al., 2014; Kaplan et al., 2016), walking and cycling refer to dissimilar behaviors that are probably affected by different urban form correlates, or by correlates that differ in their effect size (De Vries, Hopman-Rock, Bakker, Hirasing, & Van Mechelen, 2010; Aarts, Mathijssen, Van Oers, & Schuit, 2013). Failure to analyze different modes separately might result in suboptimal outcomes and less effective interventions (Giles-Corti et al., 2005; Panter et al., 2010). There is therefore a need to break the analyses down beyond a dichotomization into active and passive modes.

Finally, there is no consensus on how to incorporate urban form in mode choice models. The common practice is to model them as individual covariates (Schlossberg et al., 2006; Broberg & Sarjala, 2015). But since these variables tend to be strongly associated, multicollinearity arises, which is a problem that has thus far been disregarded (Wong et al., 2011; Lu et al., 2014). Frank et al. (2006, 2010) therefore developed a composite measure based on intra-related urban form components repeatedly constructed in studies (Kerr et al., 2006; Brownson et al., 2009; Manaugh & El-Geneidy, 2011). Another alternative is to summarize the core information of urban form variables through a few, but uncorrelated, latent variables by means of principal component analysis (PCA) (Broberg et al., 2013).

The present research addressed these three research gaps. To the best of our knowledge, it is the first comparative study of these three urban form utilizations and how they affect model interpretation. While several Anglophone studies have been published (e.g., Kerr et al., 2006; McMillan, 2007; Noland, Park, Von Hagen, & Chatman, 2014; Oliver et al., 2014; Mitra, Papaioannou, & Habib, 2016; Woldeamanuel, 2016), European studies (e.g., Müller et al., 2008; Broberg & Sarjala, 2015; Kaplan et al., 2016) are underrepresented (Lu et al., 2014). We focused on the Netherlands, a country that is highly attuned to walking and cycling (Pucher & Buehler, 2008). Although there have been a few Dutch studies (e.g., Aarts et al., 2013; Kemperman & Timmermans, 2014; Van Kann et al., 2015), exposures along travel routes were disregarded. Therefore, our objectives were:

(1) to determine the urban form correlates along children's school commuting routes affecting walking, cycling, and being chauffeured, separately; and

(2) to address the methodological issue of how to integrate urban design variables (i.e., as main effects, composite measures, or latent variables) in transport mode choice models for children's school commuting.

Our first hypothesis was that differences in urban form correlates beyond active versus passive modes are present. An understanding of which, and how, environmental correlates affect each individual mode differently is essential for more target-oriented and effective urban planning policies and physical-activity-based interventions. We assumed, secondly, that different operationalizations of urban form might not yield consistent results. Since previous studies were limited to only one approach, this would explain contradictory findings and have implications for future research.

The rest of this paper is structured as follows. We first introduce the study area and data, and then present the empirical

Figure 1. Locations of the study areas.

results. Following that, we discuss the key findings and how they translate into policy.

2. Materials and methods

2.1 Study design

The study sample comprised 97 children aged 6–11 years of six elementary schools in five cities distributed across the Netherlands (Figure 1). To minimize sociodemographic differences between the schools, schools were selected from neighborhoods that have a similar profile, although they do vary in size and population density. Each child was GPS-tracked for one regular school week between December 2008 and April 2009. Due to the mild climate and the cycling culture, a bias toward nonactive commuting can be neglected. The GPS devices (Travel recorder X, BT-Q1000X, QStarz International Co) were set up with a location sampling interval of 5 s (Kerr et al., 2011). The device was attached to an elasticated belt around the hip of each child from the morning till bedtime. For details about the study design, we referred to De Vries et al. (2010). Parents gave written consent before study inclusion. The study received approval from the Ethics Committee of the Leiden University Medical Center.

2.2 Data

2.2.1 Transport and individual and data

The raw GPS data were preprocessed in two steps. First, we applied a trip detection algorithm that separates static activity locations from trips between the residential and school location by looping chronologically through the GPS point sequence (Sterkenburg, Pierik, & De Vries, 2012). In total, 623 trips from home to school and from school to home were detected. Trips with other destinations were excluded. Second, once all school

trips had been determined, a set of decision rules and mode-specific threshold values for speed[1] were utilized for travel mode classification (Maas, Sterkenburg, de Vries, & Pierik, 2013). Of the modes, 90% were correctly classified (Sterkenburg et al., 2012), distinguishing between walking, cycling, and being chauffeured. Mode choice served as response in subsequent regressions. Public transport was disregarded, as it is an uncommon travel mode among this age cohort in the Netherlands. In-depth descriptions of GPS processing can be found elsewhere (Sterkenburg et al., 2012).

A questionnaire completed by the parents or guardians of each participating child provided personal information (e.g., age, gender). Body height and weight were collected to derive a sex- and age-specific body mass index (BMI), which was recoded to a dummy variable referring to normal weight or overweight/obesity. Personal data were attached to the GPS tracks.

2.2.2 Urban form and traffic safety exposures

Both the urban form and the traffic safety exposures were represented through a grid structure. Each city was overlaid with a regular grid with a cell resolution of 100 m. Comparisons with the actual urban morphology showed that this size was appropriate (e.g., retaining building blocks, squares). Because we doubted the relevance of exposures that are far apart, we limited the spatial context to the immediate traversed surrounding. In order to represent urban form exposures *en route*, GIS-based buffers with a width of 100 m on either side of the GPS trips were computed. Other GPS studies have used comparable buffer sizes (Badland et al., 2008). Furthermore, the influence of each cell was weighted according to its overlap with the buffers (e.g., a 50%

---

[1]The average speed of a walking trip is <10 km/h, and the maximum speed is <14 km/h. A cycling trip has an average speed of <25 km/h and a top speed of <35 km/h. Motorized trips have a maximum speed of <150 km/h.

overlap resulted in a 50% down-weighing) before the averaged exposures per variable and per trip were determined.

Guided by Brownson et al. (2009) and Wong et al. (2011), the following objectively measured urban form indices were considered. Connectivity was derived fourfold based on the street topology: the ratio between the number of cul-de-sacs, 3-way, 4-way, and >4-way intersections relative to all intersections within a cell (Schlossberg et al., 2006). Street density is represented by the total street length per cell (Panter et al., 2010). Local accessibility is the key for walking and cycling. In congruence with Helbich et al. (2016), two space-syntax-based measures were employed, namely the closeness and the betweenness indices (Cooper, 2015). While the former represents how difficult it is to reach all other locations j from a given location i, the latter refers to the business of a street segment utilizing the shortest path to go from location i to all other network locations j. Diversity is central for urban design and considered as a Shannon entropy index for land-use and building usage (Broberg & Sarjala, 2015). Land-use data stem from the Landelijk Grondgebruiksbestand Nederland 6 (LGN; Hazeu, Schuiling, Van Dorland, Oldengarm, & Gijsbertse, 2010) for the year 2007/08, while the building usage data were extracted from the cadastre (Basisregistratie Adressen en Gebouwen; BAG, 2013). Lastly, the vertical shape of the urban form was depicted through the building roughness, namely height differences between adjacent buildings (Helbich et al., 2016). This index is based on a surface model derived from Dutch laser scanning data (Waterschapshuis, 2008).

Traffic safety is an equally important determinant of mode choice (Wong et al., 2011). Three proxy variables were derived. A dummy variable indicated whether a cell intersects with a major road as indicated by Navteq street data. We controlled for the cycle path lengths relative to the total street length within a cell (Pont et al., 2009). Cycle path data were obtained from OpenStreetMap (Jokar et al., 2015). Because sidewalks are omnipresent, they were excluded. Finally, awareness of fatal and nonfatal traffic accidents is a barrier to walking and cycling (Panter et al., 2010). Cell-wise accident risk employing 2009 data on traffic accidents provided by the Netherlands Ministry of Infrastructure and the Environment was calculated by means of a kernel density estimation (i.e., higher values refer to a higher risk). Traffic volume data were not available.

2.3 Methods

2.3.1 Statistical modeling

Descriptive statistics summarize the data. Non-parametric Kruskal–Wallis tests were used for continuous variables, and Fisher tests were used for categorical variables to investigate differences across the transport modes statistically. Spearman's correlation coefficients were used to test the associations among the urban form variables.

To test the associations between children's transport mode choice (i.e., walking, cycling, being chauffeured) and the covariates, we estimated multinomial logit regression models (Belitz et al., 2015). The models were set up as structured additive regressions with two random effects and fitted through restricted maximum likelihood (Umlauf, Adler, Kneib, Lang, & Zeileis, 2015). The first random effect was needed because individual trips were repeatedly taken by the same child, while the second one was required because the children were nested within cities. Neglecting such correlations challenges the validity of statistical inference (Gelman & Hill, 2007). In all models, being chauffeured served as baseline (reference group).

To obtain information about individual effects, univariate models were fitted first. Variables significant at the $p < 0.1$ level were used to develop multivariate models. In total, five multivariate models were estimated. While the base model is without urban form variables, models 2–5 include urban form differently (e.g., as main effect, PCs, composite measure). The models' goodness-of-fits were compared with the Akaike information criterion (AIC), and a suitable random-effects structure was determined. The lower the AIC score, the better the model fits.

2.3.2 Principal component analysis and composite measure

Modeling urban form variables as main effects makes interpretation easier, but correlations arising among these covariates make individual effects hardly assignable and challenges statistical inference (Noland et al., 2014; Lu et al., 2014). To overcome this, either PCA (Jolliffe, 2002) can be performed or an index can be constructed.

PCA transforms the original urban form variables into latent ones depicting the overall data variability through a linear combination of the (scaled) variables represented by a few, uncorrelated PCs. The first PC accounts for most of the variability in the data, while the succeeding PCs capture the remaining but declining variability. As PCA extracts an equal number of PCs as available input variables, the determination of a suitable number of PCs is crucial (Jolliffe, 2002). Screeplots in combination with the Kaiser criterion are useful. The former plots the number of components against the eigenvalues, and the "elbow" refers to an appropriate number of PCs, while the latter signifies PCs with eigenvalues greater than 1 as suitable. To ease the PCA interpretation, the bi-plot is supportive (Jolliffe, 2002). The loadings (i.e., arrows) show the direction of each PC, expressing the associations to the original variables, whereas the scores are the new data points projected on the new coordinates for each PC. Once identified, these PCs enter the mode choice models.

Alternatively, Frank et al. (2006, 2010) proposed a composite measure of urban form. This index is based on intersection density, land-use diversity, etc. and sums the z-scores of the involved variables. Despite several construction possibilities (Kerr et al., 2006; Clark, Scott, & Yiannakoulias, 2014), a comparison by Manaugh and El-Geneidy (2011) showed only minor differences for travel behavior modeling, highlighting their good performance in summarizing the dimensions of urban design.

3. Results

Of the trips, 64% were done on foot, 27% by bike, and 9% by automobile. Key descriptive statistics for each variable stratified by transport mode are shown in Table 1. With $p < 0.001$, both the Kruskal–Wallis and the Fisher tests refer to significant differences across the modes for several variables, including age,

Table 1. Descriptive statistics.

	Walking ($n = 400$)	Cycling ($n = 170$)	Motorized ($n = 53$)	p-Value
Personal characteristics				
Age (median [IQR])	10.00 [8.00, 10.00]	9.00 [8.00, 10.00]	7.00 [6.00, 9.00]	<0.001
Sex (0 = female, 1 = male) (%)	262/138 (65.5/34.5)	94/76 (55.3/44.7)	21/32 (39.6/60.4)	<0.001
BMI (0 = normal weight, 1 = overweighted) (%)	132 (33.0)	41 (24.1)	6 (11.3)	0.001
Trip characteristics				
Direction (0 = home–school, 1 = school–home	184/216 (46.0/54.0)	95/75 (55.9/44.1)	28/25 (52.8/47.2)	0.084
Distance in meters (logged) (median [IQR])	5.36 [3.51, 5.90]	5.79 [5.49, 6.45]	6.81 [6.47, 7.06]	<0.001
Traffic safety				
Major roads (median [IQR])	0.00 [0.00, 0.01]	0.00 [0.00, 0.01]	0.00 [0.00, 0.12]	0.001
Cycle path (median [IQR])	0.76 [0.02, 9.69]	2.33 [0.11, 24.90]	1.06 [0.00, 10.80]	0.005
Accident risk (median [IQR])	0.09 [0.08, 0.20]	0.12 [0.09, 0.24]	0.23 [0.07, 0.26]	<0.001
Urban form				
Cul-de-sac (median [IQR])	2.83 [0.00, 7.01]	5.49 [2.44, 8.36]	7.10 [2.55, 10.20]	<0.001
3-way crossings (median [IQR])	31.05 [20.80, 35.20]	29.55 [22.10, 34.85]	34.00 [26.50, 36.30]	0.128
4-way crossings (median [IQR])	4.79 [3.81, 7.72]	6.65 [3.59, 9.71]	5.25 [2.68, 8.04]	0.032
>4-way crossings (median [IQR])	0.15 [0.00, 1.37]	0.00 [0.00, 0.77]	0.00 [0.00, 0.91]	0.001
Land-use diversity (median [IQR])	0.04 [0.02, 0.13]	0.13 [0.04, 0.19]	0.15 [0.05, 0.18]	<0.001
Building diversity (median [IQR])	0.13 [0.04, 0.20]	0.16 [0.12, 0.25]	0.14 [0.12, 0.17]	<0.001
Roughness index (median [IQR])	0.14 [0.11, 0.19]	0.18 [0.12, 0.26]	0.25 [0.22, 0.32]	<0.001
Closeness index (median [IQR])	60.80 [53.70, 67.88]	67.75 [62.00, 72.00]	74.70 [70.90, 78.40]	<0.001
Betweenness index (median [IQR])	24.20 [16.20, 27.20]	22.15 [13.93, 26.28]	15.40 [13.80, 17.50]	<0.001
Street density (median [IQR])	106.00 [87.90, 120.00]	114.50 [91.25, 127.75]	116.00 [105.00, 125.00]	0.001

IQR, interquartile range.
p-values are based on the Kruskal–Wallis test (continuous variables) or the Fisher test (categorical variables).

gender, and trip distance, and several urban form variables. For example, the median age at which younger children are more prone to be chauffeured than older ones varies. A motorized mode is more frequently used for longer travel distances compared to walking and cycling. Some of the correlations among the urban form covariates are critical. For example, the roughness and closeness indices are, with a value of 0.720 ($p < 0.001$), rather strongly associated.

To analyze individual associations between mode choice and the covariates, univariate models were estimated. AIC scores

Table 2. Results of the univariate multinomial mixed models.

	Walking Coef.	p-Value	Cycling Coef.	p-Value	AIC
Personal characteristics					
Age	0.408	0.023*	0.342	0.076#	798.143
Sex	−0.939	0.062#	−0.826	0.115	794.064
BMI (n.s.)	0.876	0.149	0.377	0.553	797.606
Trip characteristics					
Direction (n.s.)	0.515	0.134	−0.052	0.886	791.707
Distance (logged)	−0.853	0.000***	−1.017	0.000***	780.841
Traffic safety					
Major roads	−10.551	0.001**	−9.038	0.016*	791.059
Cycle path (n.s.)	−0.008	0.671	−0.009	0.620	798.088
Accident risk (n.s.)	0.365	0.827	0.475	0.784	797.961
Urban form					
Cul-de-sac (n.s.)	0.011	0.811	−0.016	0.730	798.183
3-way crossings	−0.073	0.012*	−0.062	0.044*	794.852
4-way crossings (n.s.)	0.046	0.497	0.072	0.313	798.702
>4-way crossings	0.226	0.241	0.475	0.019*	794.067
Land-use diversity	−11.604	0.000***	−8.305	0.009**	786.933
Building diversity (n.s.)	2.566	0.611	0.314	0.951	797.748
Roughness index	−8.920	0.000***	−9.320	0.000***	777.661
Closeness index	−0.121	0.000***	−0.116	0.000***	779.031
Betweenness index	0.093	0.035*	0.146	0.001**	793.299
Street density	−0.038	0.000***	−0.029	0.012*	792.459

Being chauffeured served as reference category.
n.s. refers to non-significant variables.
AIC, Akaike information criterion.
Signif. codes: <0.001 '***', <0.01 '**', <0.05 '*', <0.1 '#'.

consistently refer to better model performances using a person-based and a city-based random effect, instead of employing only a person-based random effect. Summary statistics of these univariate models are given in Table 2. Insignificant variables with $p > 0.1$ (i.e., BMI, trip direction, bike path, accident risk, cul-de-sac, 4-way crossings, building diversity) were removed from subsequent analyses.

The significant urban form variables of the univariate models were then used to derive the composite measure and to conduct PCA. The first quantile of the composite index is −5.639 and the third is 5.138, with a mean of around 0 (S.D. = 7). The Kaiser criterion (eigenvalues >1) shown in Figure 2 (left panel) suggests that the first three PCs explain in total approximately 80% of the total variance (PC1: 33%, PC2: 32%, PC3: 15%). Collinearity is no longer present among the PCs. The bi-plot (Figure 2, right panel) shows both grouping effects of the variables and similar loadings. PC1 is mainly characterized by negative loadings, while PC2 shows a balance between negative and positive loadings. PC1 shows the tendency to encompass connectivity, density, and diversity variables, while PC2 comprises rather more accessibility variables. Still, the interpretation remains demanding. Several urban form input variables are highly correlated, as indicated by the angle between two arrows (e.g., the closeness and building roughness indices).

Table 3 shows the estimation results of the five multivariate models equipped with two random effects, as suggested by the AIC. With an AIC score of 771, the PCA-based model 4 has the best fit, followed by model 2 (ΔAIC = 4.4) and model 3 (ΔAIC = 7.4), while model 5 with the composite measure performs poorly and is only slightly better (ΔAIC = 5.8) than the base model.

Even the input variables are significant in the univariate models, whereas gender and major roads turned out to be insignificant in the multivariate base model. Age is significantly positively related to walking ($\beta = 0.441$, $p < 0.010$) and only weakly to cycling ($\beta = 0.394$, $p < 0.100$). Trip distance has a negative effect on walking ($\beta = -0.694$, $p < 0.001$) and cycling

($\beta = -0.905$, $p < 0.001$) relative to the automobile. In the best performing model (4), age remains a weakly significant covariate for walking ($\beta = 0.321$, $p < 0.010$) but not for cycling. This contrasts with model 5, which incorporates the composite urban form index, where age is significantly positively associated with both walking and cycling. No significant effect of major roads emerged in models 4 and 5. After extending the base model with latent urban form variables (i.e., PC1-3), it turns out that distance is no longer a discriminating factor for mode choice. Moreover, in model 4, PC1 and PC2 are significantly related to mode choice, though not PC3. However, distance remains strongly significant ($\beta_{walking} = -0.635$, $p < 0.002$; $\beta_{cycling} = -0.882$, $p < 0.001$) when the composite index is applied (model 5). In both model 4 and model 5, the challenge concerning the interpretation of the urban form associations persists. Thus, model 2 considers urban forms as an individual variable. Age is positively related to mode choice ($\beta_{walking} = 0.520$, $p < 0.005$; $\beta_{cycling} = 0.391$, $p < 0.093$), while gender is related only to a minor extent to walking ($\beta = -0.881$, $p < 0.091$). Like in model 4, distance is not significant, and the univariate effect of main roads disappears. For walking only, 3-way intersections ($\beta = -0.094$, $p < 0.064$), >4-way intersections ($\beta = -0.586$, $p < 0.023$), and land-use diversity ($\beta = -11.970$, $p < 0.021$) show a negative association. Furthermore, the two accessibility measures are statistically significant. While closeness index shows a consistent negative association with walking ($\beta = -0.102$, $p < 0.034$) and cycling ($\beta = -0.102$, $p < 0.071$), the betweenness index indicates a positive relationship for walking ($\beta = 0.183$, $p < 0.001$) and cycling ($\beta = 0.128$, $p < 0.070$). Despite being significant in the univariate models, neither building roughness nor street density shows any significance in the multivariate model.

4. Discussions

In the present study, we examined how urban form correlates are related to children's commuting mode choice. By utilizing GPS tracking, modeling associations for transport modes separately, and comparing how different urban form operationalizations (i.e., individual variables, composite measures, latent variables) alter model interpretation, we have contributed to the existing literature.

4.1 Descriptive statistics

Our descriptive statistics showed a large share of walking and cycling, whereas in North America and Australia, chauffeuring is pervasive (Carver et al., 2013). In the Netherlands, active commuting (64% walking and 27% cycling trips; total 91%) is more prevalent than that in, for example, Norway (75%) (Børrestad, Andersen, & Bere, 2011) or Belgium (59%) (D'Haese, De Meester, De Bourdeaudhuij, Deforche, & Cardon, 2011). Our total percentage of active commuting trips than that reported for the hilly Limburg area of the Netherlands (64%) (Van Kann et al., 2015). That the Netherlands has an overwhelming share of active travelers is not surprising, considering that the population is highly disposed toward walking and cycling, the country has a mild climate and a high-standard walking and cycling infrastructure, and school are physically well-integrated in neighborhoods (Pucher & Buehler, 2008; Helbich, Böcker, & Dijst, 2014).

4.2 Principal findings concerning the individual effects

Model 2 showed that a limited number of variables are associated with children's mode choice, confirming the results of Oliver et al. (2014). Concerning children's individual characteristics, a persistent finding across the models was that age is positively related to walking and cycling, with a distinct effect for the former. This confirms the findings of Aarts et al. (2013) as well as Goeverden and De Boer (2013) that older children are more likely to walk or cycle. This behavior can be explained through improved cognitive capabilities, which result in more independent mobility (Mitra, 2013). The literature is inconclusive about gender effects on mode choice (Faulkner et al., 2009). Although univariate models indicated that mode choice varies across sex, analogous to Larsen et al. (2009), no evidence was found in multivariate models. While supporting Van Goeverden and De Boer (2013) for the Netherlands and McDonald (2008) for the United States that gender is not significantly related to walking, this result contradicts the position of Børrestad et al. (2011), who showed that Norwegian girls are more likely to walk and less likely to cycle compared to boys.

In the base model, the trip distance stands out. Without adjusting for urban form, distance is, as expected, a strong discriminating factor, as is the case in the literature (Bringolf-Isler et al., 2008; Müller et al., 2008; Larsen, Gilliland, & Hess, 2012; Kemperman & Timmermans, 2014). This is intuitive, since longer distances require pronounced physical effort, which makes walking and cycling less attractive (Aarts et al., 2013; Panter, Corder, Griffin, Jones, & Van Sluijs, 2013). However, this conclusion must be reversed when the model is adjusted for urban form, modeling the variables either directly or through PCs. It seems that distance-related effects get absorbed by the built environment (Helbich et al., 2016). This does not occur when the composite measure is considered instead. It could be speculated that the index represents urban form less holistically compared to the PCs, as evidenced by the weaker model performance. Or perhaps the importance of the involved variables is not reflected appropriately (e.g., Frank et al. (2010) weighed the involved variables differently). As reviewed by Pont et al. (2009), others also report insignificant distance effects. Trip direction appeared insignificant, which confirms the suggestion by Larsen et al. (2009) that travel decisions are usually made for the entire day.

As for transport safety and urban form correlates, only a few were associated with mode choice, as Panter et al. (2013) found in their study in Norfolk (United Kingdom). Neither walking nor cycling was significantly related to traffic safety, as also found by Schlossberg et al. (2006). This contradicts Van Kann et al. (2015), who reported that traffic safety (e.g., cycling infrastructure, crossing major roads) is a major issue for walking and cycling, but supports Ewing et al. (2004) and Aart et al. (2013), who showed that bike paths are insignificant. However, these results infer not that traffic safety is irrelevant, but that in the Netherlands traffic hazards have been minimized by nearly omnipresent sidewalks, bike lanes, etc. Or perhaps that the environment as perceived by the parents represents safety

issues more realistically than the utilized GIS measures (Mitra, Faulkner, Buliung, & Stone, 2014).

Whereas functional building usage seems irrelevant for mode choice, the opposite is true for land-use diversity where lower index scores only foster walking. Although several studies found null associations (Wong et al., 2011), we found a significantly negative association for walking. This could be because the schools in our study are not located in inner cities, which have pronounced land-use diversity. Aarts et al. (2013) argued that building height differences might encourage active modes, but we are not able to support this hypothesis.

In keeping with previous studies (Schlossberg et al., 2006; Panter et al., 2010; Larsen et al., 2012), street connectivity seems essential. Model 2 suggests that 3-way and >4-way intersections are inversely and significantly associated with walking but not with cycling. This means, for example, that better-connected street networks (i.e., >4-way intersections) decrease walking relative to the automobile. At the first glance, this relationship appears counterintuitive, but well-connected streets also increase the risk of traffic accidents due to crossings. Wong et al. (2011) also found that other intersection (e.g., cul-de-sac) variables are insignificant. Cross-comparisons with other study sites are not feasible for the space-syntax-based accessibility measures because these have not previously been considered. For walking and cycling, the closeness shows a negative association, meaning that a less-accessible street network reduces the likelihood for both active modes relative to the automobile usage. A reverse interpretation applies to the betweenness index, where well-accessible streets promote walking and cycling (Helbich et al., 2016). These results are rational, since in the Netherlands accessibility is a major locational factor in selecting a primary school (Herweijer & Vogels, 2004).

4.3 Principal findings concerning urban form operationalizations

Although intuitive to do, entering the urban form correlates as main effects might create multicollinearity problems (Broberg et al., 2013; Noland et al., 2014; Lu et al., 2014). However, PCA and composite measures present a way to overcome this challenge (Manaugh & El-Geneidy, 2011; Clark et al., 2014).

While reducing collinearity, both PCA and composite measures impede the direct interpretation of how environmental variables are associated with transport behavior. The former is also challenged by the selection of a suitable number of PCs, even though several criteria (e.g., the Kaiser criterion) alleviate this decision (Jolliffe, 2002). These criteria usually suggest more than only the first PC (e.g., Broberg et al., 2013 extracted five PCs). For both PCA and composite measures, it remains unclear how each individual factor is related to children's mode choice (Noland et al., 2014). However, such knowledge might be crucial to implement environment-based interventions and set up policies that promote physical activity. To make it easier to gain an in-depth understanding of how the derived latent variables are formed, we suggest using visualization tools such as the bi-plot. Although Clark et al. (2014) argued for the simplicity of composite measures, the risk that positive and negative values might counteract each other remains.

As shown, an inappropriate number of PCs deteriorate the goodness-of-fit (model 3 vs. model 5), which calls for careful selection. Comparing different composite measures, Manaugh and El-Geneidy (2011) ascribe them, in general, a good performance in explaining travel behavior. Our results show otherwise, as indicated by the model performances. A worse AIC score is achieved by model 5 compared to model 4. This might be interpreted as meaning that the composite measure does not effectively represent the entire complexity of urban form better than the representation based on statistical dimension reducing techniques (Broberg et al., 2013). Because the "true" associations are unknown, we cannot recommend one specific method as the gold standard but, based on our findings, we lean toward PCA.

4.4 Strengths and limitations

The present study contributes to the limited number of GPS-based case studies on school commuting. *En route* urban form exposures represent the environmental context more realistically than Euclidean distances or shortest path analyses, assuming that children optimize their routes (Harrison et al., 2014). Besides, the route recall bias introduced by self-administered travel diaries is eliminated. Whereas the majority of studies look at walking and cycling as a single category (Bringolf-Isler et al., 2008; Larson et al., 2009; Broberg et al., 2013), we estimated urban form associations separately for individual transport modes. Another key strength is that urban form variables were operationalized on a spatial micro level, considering the local morphology on both the horizontal and the vertical level. This extension toward the third dimension is unique, as is the application of space-syntax-based urban form indicators. Because the Netherlands was the study area, this investigation advances the literature specifically targeting a walking- and cycling-oriented country (Van Goeverden & de Boer, 2013), which complements the automobile-oriented North America-centric body of knowledge (Pont et al., 2009). Finally, the present research was the first to compare different urban form operationalizations.

Although it breaks new ground, this study has limitations. Though useful for hypotheses generation, the research design is cross-sectional (Lu et al., 2014; Fishman et al., 2015). We recommend that future studies be longitudinal and consider changes of the covariates over time (Mitra et al., 2016). Parent's attitudes toward their residential location choice were not available and self-selection cannot be ruled out (Mokhtarian & Cao, 2008). Yet, carefully following Badland et al. (2008) and Wong et al. (2011), we cannot obviate an ideal context representation, calling for sensitivity studies. We focused on school commuting trips, but Kemperman and Timmermans (2014) showed that other trip purposes are affected differently by urban form. The sampling was restricted to children living in the school neighborhood. Even though long commuting distances are uncommon for this age group, it might have an influence on mode choices, especially public transport. Although the influence of socioeconomics on the household level (e.g., income, automobile ownership) is inconclusive (McDonald, 2008; Van Goeverden & de Boer, 2013), our models remained unadjusted due to missing data and did not acknowledge interdependence among

transport mode alternatives (Ermagun & Levinson, 2017). Lastly, our dataset solely relied on objective indicators, although there is no doubt that parental perceptions influence children's mobility (Yeung et al., 2008; Mitra, 2013). It would be informative to enrich our data in that respect, and to consider neighborhood cohesion, stranger danger, etc. (Chillon et al., 2014).

4.5 *Implications for research and policy*

The results are valuable for both academia and policymakers. We suggest that conclusions drawn from mode choice models depend on how urban form correlates are integrated. Although Adams et al. (2014) introduced common ground to derive such measures, integrating these indices in statistical models remains a challenge. Careful attention must be paid to correlations among urban form proxy variables, and thorough screening for multicollinearity among variables is recommended. We encourage researchers to go beyond only one representation, since findings might change with changes in urban form operationalization. The number of PCs should be based on one of the well-established criteria to avoid oversimplification and to have sufficient data-inherent information. The research community should also recognize the availability of models that are robust against multicollinearity (e.g., elastic net; Zou & Hastie, 2005) and start to use them (Helbich et al., 2016).

Most policies are broadly designed to increase active school commuting, and they combine walking and cycling into a single category. Such "one-fits-all" policies may be inefficient, or interventions may fail. The present research will help planning authorities and public health strategists to understand the complex way in which urban environmental correlates influence transport behavior. Our findings call for interventions that are targeted at either walking or cycling, rather than both: Only interventions developed for a specific transport mode will yield significant and population-wide health benefits.

5. Conclusions

This analysis evidences not only that urban form correlates along the commuting routes vary across children's transport modes, but also that the magnitudes of the estimated coefficients differ in size. To optimize the effectiveness of planning and health interventions, it is key for policymakers to develop mode-specific actions dedicated to either walking or cycling, and not both. Secondly, not only were inconsistent results found for different urban form representations, but also the associations among the children's personal characteristics and trip distance changed. As the models' goodness-of-fits indicate distinct differences across urban form operationalizations, a well-selected number of PCs, rather than an urban form composite measure, should be used. Sensitivity analyses are thus recommended.

Acknowledgments

Funding

References

Aarts, M. J., Mathijssen, J. J., Van Oers, J. A., & Schuit, A. J. (2013). Associations between environmental characteristics and active commuting to school among children: A cross-sectional study. *International Journal of Behavioral Medicine*, *20*(4), 538–555.

Adams, M. A., Frank, L. D., Schipperijn, J., Smith, G., Chapman, J., Christiansen, L. B., Coffee, N., Salvo, D., Toit, L. d., Dygrýn, J., Ferreira Hino, A. A., Lai, P.-C., Mavoa, S., Pinzón, J. D., Van de Weghe, N., Cerin, E., Davey, R., Macfarlane, D., Owen, N., & Sallis, J. F. (2014). International variation in neighborhood walkability, transit, and recreation environments using geographic information systems: The IPEN adult study. *International Journal of Health Geographics*, *13*(1), 1.

Badland, H. M., Schofield, G. M., & Garrett, N. (2008). Travel behavior and objectively measured urban design variables: Associations for adults traveling to work. *Health & Place*, *14*(1), 85–95.

BAG. (2013). *Kadaster*. Retrieved 21 December 2013, from https://www.kadaster.nl/bag

Bauman, A. E., Reis, R. S., Sallis, J. F., Wells, J. C., Loos, R. J., Martin, B. W., & Lancet Physical Activity Series Working Group. (2012). Correlates of physical activity: Why are some people physically active and others not? *The Lancet*, *380*(9838), 258–271.

Belitz, C., Brezger, A., Klein, N., Kneib, T., Lang, S., & Umlauf, N. (2015). BayesX—Software for Bayesian inference in structured additive regression models. Retrieved from http://www.bayesx.org

Böcker, L., Dijst, M., Farber, J, & Helbich, M. (2015). En-route weather and place valuations for different transport mode users. *Journal of Transport Geography*, *47*, 128–138.

Børrestad, L. A., Andersen, L. B., & Bere, E. (2011). Seasonal and socio-demographic determinants of school commuting. *Preventive Medicine*, *52*(2), 133–135.

Broberg, A., Salminen, S., & Kyttä, M. (2013). Physical environmental characteristics promoting independent and active transport to children's meaningful places. *Applied Geography*, *38*, 43–52.

Broberg, A., & Sarjala, S. (2015). School travel mode choice and the characteristics of the urban built environment: the case of Helsinki, Finland. *Transport Policy*, *37*, 1–10.

Brownson, R. C., Hoehner, C. M., Day, K., Forsyth, A., & Sallis, J. F. (2009). Measuring the built environment for physical activity: State of the science. *American Journal of Preventive Medicine*, *36*(4), S99–S123.

Bringolf-Isler, B., Grize, L., Mäder, U., Ruch, N., Sennhauser, F. H., & Braun-Fahrländer, C. (2008). Personal and environmental factors associated with active commuting to school in Switzerland. *Preventive Medicine*, *46*(1), 67–73.

Carver, A., Timperio, A., & Crawford, D. (2013). Parental chauffeurs: What drives their transport choice? *Journal of Transport Geography*, *26*, 72–77.

Cervero, R., & Kockelman, K. (1997). Travel demand and the 3Ds: Density, diversity, and design. *Transportation Research Part D: Transport and Environment*, *2*(3), 199–219.

Chillón, P., Hales, D., Vaughn, A., Gizlice, Z., Ni, A., & Ward, D. S. (2014). A cross-sectional study of demographic, environmental and parental barriers to active school travel among children in the United States. *International Journal of Behavioral Nutrition and Physical Activity*, *11*(1), 61.

Clark, A. F., Scott, D. M., & Yiannakoulias, N. (2014). Examining the relationship between active travel, weather, and the built environment: