Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 00 (2017) 000–000

Transportation Research Procedia

www.elsevier.com/locate/procedia

World Conference on Transport Research - WCTR 2016 Shanghai. 10-15 July 2016

Human factors of pedestrian walking and crossing behaviour

Eleonora Papadimitriou[a,]*, Sylvain Lassarre[b], George Yannis[a]

[a]*Department of Transportation Planning and Enginnering, NTUA, 5 Heroon Polytechniou str., GR-15773, Athens, Greece*
[b] *IFSTTAR - Institut français des sciences et technologies des transports, de l'aménagement et des réseaux, 14-20 Boulevard Newton, Cité Descartes, Champs sur Marne, F-77447 Marne la Vallée Cedex 2, France*

Abstract

Human factors related to pedestrians have received somewhat less attention in the literature compared to other road users, although it is often underlined that road and traffic factors appear to explain only a small part of pedestrian walking and crossing behaviour in urban areas. The understanding of pedestrian behaviour in urban areas may assist in the improved design and planning of the road and traffic environment, and consequently to the improvement of pedestrian comfort and safety. The objective of this research is the exploration of human factors of pedestrian walking and crossing behaviour in urban areas. More specifically, this research aims to capture and analyse key components affecting pedestrian walking and crossing behaviour, namely the pedestrians' attitudes, perceptions, motivations, behaviour and habits.

A questionnaire was designed aiming to capture key human factors of pedestrian walking and crossing behaviour including their mobility characteristics and travel motivations, their risk perception and their value of time, their attitudes towards walking and related preferences, their walking and crossing behaviour and compliance to traffic rules, their self-assessment, their opinion on drivers etc. The questionnaire included 54 questions and the responses were given on a 5-point Likert scale (e.g. from "strongly agree" to "strongly disagree", from "never" to "always"), plus some basic questions on demographics. The questionnaire was filled by 75 survey young and middle-aged participants, out of which 40 were males. A thorough descriptive analysis of the questionnaire data was carried out, in order to identify main trends and patterns. A principal component analysis of the data was then implemented, in order to identify underlying factors ("components") of pedestrian walking and crossing behaviour.

The descriptive analysis of the questionnaire responses revealed that most pedestrians have positive attitudes, preferences and behaviours (e.g. risk-conscious and compliant); nevertheless, there is a non-negligible proportion of pedestrians who have negative attitudes and are willing to make dangerous actions (e.g. cross diagonally or at mid-block). A PCA results suggest that there are three dimensions of human factors of pedestrian behaviour: the first two concern their risk perception and risk taking (one reflecting risky attitudes behaviours and the other one reflecting conservative attitudes and behaviours) and the third one concerns walking motivations. There are also two groups of pedestrians identified by a cluster analysis over the dimensions scores: "positive and motivated" vs. "negative and unmotivated" pedestrians.

* Corresponding author. Tel.: +30 210 7721380; fax: +30 210 7721454.
E-mail address: dpavlou@central.ntua.gr

2214-241X © 2017 The Authors. Published by Elsevier B.V.
Peer-review under responsibility of WORLD CONFERENCE ON TRANSPORT RESEARCH SOCIETY.

Peer-review under responsibility of WORLD CONFERENCE ON TRANSPORT RESEARCH SOCIETY.

Keywords: pedestrian behaviour; human factors; principal component analysis, cluster analysis.

1. Background and objectives

The literature on human factors and road user behaviour in road and transport design is extensive (e.g. Fuller & Santos, 2002). Human factors related to pedestrians have received somewhat less attention in the literature compared to other road users, although it is often underlined that road and traffic factors alone may explain only a small part of pedestrian walking and crossing behaviour in urban areas (Papadimitriou, 2012). The understanding of pedestrian behaviour in urban areas may assist in the improved design and planning of the road and traffic environment, and consequently to the improvement of pedestrian comfort and safety. On the other hand, there are several studies dealing with human factors in pedestrian road crossing decisions, however these factors are examined alone, outside the context of the road and traffic environment. As a result, these studies provide useful insight on the behavioural and psychological aspects of pedestrian decisions, but have little applicability in terms of describing the crossing behaviour in urban areas (Papadimitriou et al. 2009).

The objective of this research is the in-depth analysis of human factors of pedestrian walking and crossing behaviour in urban areas. More specifically, this research aims to capture and analyse key components affecting pedestrian walking and crossing behaviour, namely the pedestrians attitudes, perceptions, motivations, behaviour and habits, and identify 'profiles' of pedestrians on the basis of these human factors.

1.1. Literature review

There are numerous studies dealing with human factors of pedestrians' crossing behaviour, using formal tools such as questionnaires or in-depth interviews. Hine (1996) used in-depth interviews to identify pedestrians' perception as regards difficulty to cross and assessment of traffic conditions and crossing facilities in the centre of Edinburgh.

Evans and Norman (1998) developed hierarchical regression models for road crossing behaviour, by means of completed questionnaires which included scenarios of three specific potentially dangerous road crossing behaviours. Pedestrians stated crossing behaviour was then modelled in relation to measures of attitude, subjective norm, perceived behavioural control, self-identity and intention. Yagil (2000) proposed multivariate regression models for the self-reported frequency of unsafe crossings in relation to beliefs regarding the consequences of the behaviour, instrumental and normative motives for compliance with safety rules, and situational factors.

Diaz (2002) developed a structural equations model for explaining pedestrian risk-taking behaviour on the basis of attitude, subjective norm, perceived control, behavioural intention and reported violations, errors and lapses. Self-reported crossing behaviour data from pedestrians in the city of Santiago was used on that purpose. Holland and Hill (2007) tested for age and gender differences in road crossing decisions within a theory of planned behaviour analysis including intention, situation and risk perception effects. Oxley et al. (1997) examined the crossing behaviour of elderly pedestrians at mid-block locations by measuring a number of indicators such as kerb delay, gap acceptance, crossing time, time-of-arrival, minimum safety margin and crossing style (non-interactive vs. interactive). Results showed that elderly pedestrians present increased kerb delay, and accept larger gaps; however they also frequently adopt unsafe interactive crossing styles.

A related study (Bernhoft and Carstensen, 2008) revealed that older pedestrians appreciate sidewalks and crossing facilities much more than younger pedestrians. Rosenbloom et al (2008) used a similar method to examine the crossing behaviour of children and found that not looking was the most prevalent unsafe behaviour, followed by the combination of not looking and not stopping, and not stopping before crossing. They also found that children accompanied by an adult committed more unsafe behaviours, especially when not holding hands with the adult.

Sisiopiku & Akin (2003) define spatial crossing compliance rate (SCCR) (i.e. how pedestrians use crosswalks with respect to the location) and temporal crossing compliance rate (TCCR) (i.e. how pedestrians use red signal at traffic controlled locations). The study used both questionnaire and field survey data to compare the perceived and observed compliance rates and concluded that unsignalised mid-block crosswalks were the treatment of preference to pedestrians with a high compliance rate. It was also found that the crosswalk location, relative to the origin and destination of pedestrians, was the most influential decision factor for pedestrians deciding to cross at a designated location.

Recently, within the SARTRE-4 European survey on road users' attitudes, perceptions and behaviour, a dedicated questionnaire was addressed to pedestrians from 19 countries. The statistical analysis that followed (Papadimitriou et al. 2013), revealed seven components of pedestrian attitudes and behaviour (formed on the basis of 54 questionnaire elements), namely: 'satisfaction with pedestrian environment', 'attitudes towards penalties', 'attitudes towards electronic in-vehicle devices', 'attitudes towards speed limitations and surveillance', 'pedestrian behaviour and distraction', 'attitudes towards pedestrian safety design', 'annoyance with other road users' and 'lack of accessibility'. Furthermore, pedestrians were clustered in three groups: "positive attitudes and positive behaviour", "negative attitudes and negative behaviour", and "mixed attitudes and positive behaviour", with significant variation over gender, age groups and countries.

Granié et al. (2013) developed and tested a questionnaire for the self-reporting behaviour of pedestrians. Out of 47 elements tested, 20 were found to be most important for assessing pedestrian behaviour, within 4 components, namely: "transgression", "lapses", "aggressive behaviour" and "positive behaviour".

The existing studies on human factors of pedestrian behaviour are summarized in Table 1.

Table 1. Overview of studies on human factors of pedestrian behaviour

		Problem			Theory		Models			Variables			Data	
Author	Year	crossing decision	perceptions etc.	compliance	Planned behaviour	Other	Regression, GLM	Factor analysis	Other	individual	roadway	traffic	observational data	self-reported
Hine	1996	●	●			●			●		●	●		●
Evans and Norman	1998	●	●		●		●			●				●
Yagil	2000	●		●	●		●			●				●
Diaz	2002	●	●		●			●		●				●
Holland and Hill	2007	●			●		●			●				●
Oxley et al.	1997	●	●			●			●	●			●	
Berhhoft and Carstensen	2008	●		●		●	●			●	●			●
Rosenbloom et al.	2008	●		●		●	●			●			●	
Sisiopiku & Akin	2003	●		●		●	●			●	●	●	●	●
Papadimitriou et al.	2013		●	●		●		●		●	●	●		●
Granié et al.	2013		●	●		●		●		●				●

1.2. Research hypotheses

On the basis of the literature review, the human factors to be examined in the present research were defined, and specific research hypotheses to be tested were formulated, as follows:

- Demographics
  - Age: Younger pedestrians are more risk-taking and less compliant to traffic rules related to road crossing
  - Gender: Male pedestrians are more risk-taking.
  - Income: low income, perceived social inequality and the lack of alternatives to walking may lead pedestrians to more aggressive and less compliant behaviour.
- Travel motivations

  - Walking frequency and distance travelled: a positive relationship between walking frequency / distance travelled and crossing behaviour is assumed
  - Walking purposes: pedestrians walking for health / recreation purposes are likely to be less risk-taking and more safety conscious
- Risk perception and value of time
  - The risk minimizer: minimizes the number of crossings and increases the length of the path in order to avoid vehicle / pedestrian interaction;
  - The delay minimiser: maximizes the number of crossings in order to reduce the length of the pedestrian path;
- Traffic behaviour and compliance
  - Crossing behaviour: pedestrians self-reported crossing behaviour is similar to their observed behaviour.
  - Traffic law compliance: more compliant and less risk-taking pedestrians are less likely to cross outside designated locations;
- Interaction with other road users
  - Imitation and leader / follower effects: some pedestrians may 'follow' the crossing choices of others, while others may prompt their company to a specific behaviour
  - Opinion towards drivers: Pedestrians with negative opinion on drivers are more likely to be careful and compliant

2. Data collection

In order to test the above research hypotheses, a questionnaire was designed and implemented. The questionnaire was eventually created as a list of items to be rated on the basis of Likert scales expressing always/never or agree/disagree scales. The self-reported behavioural questionnaire of Granié et al. (2013) was used as a basis: a selection of questions on behaviour and compliance was carried out, complemented with elements on perceptions, attitudes, beliefs, motivations etc. from other published questionnaires (Evans & Norman, 1998; Bernhoft & Carstensen, 2008; Yagil, 2000; Sisiopiku & Akin, 2003). The synthesis was completed with the introduction of some additional specific elements that were of particular interest in this research.

The questionnaire includes 6 sections, namely: (a) demographics, (b) mobility and travel motivations, (c) attitudes, perceptions and preferences, (d) self-assessment and identity, (e) behaviour, compliance and risk taking, and (f) opinion on drivers, and it is presented in Table 2. The questionnaire was filled by 75 individuals in the period July - December 2013; 53% of the participants were males, 50% of the participants were 18-24 years old, 27% were 25-34, 20% were 35-45 and 3% were >45 years old. The majority of the participants are frequent pedestrians, as more than 50% reported daily frequency of pedestrian trips. However, they also use both private cars and public transport.

Table 2. Questionnaire design

A	**Demographics**
A_i.	Gender (female / male)
A_ii.	Age group (18-25, 25-35, 35-45, 45-55, 55-65)
A_iii.	Annual income (Euros) (<5000, 5000-10.000, 10.000-20.000, >20.000)
B	**How many times per week do you travel by each one of the following modes*:**
B1_i	Public transport (metro, bus, trolley bus, tramway)
B1_ii	Pedestrian
B1_iii	Passenger car (driver or passenger)
	Last week, how many kilometres did you travel by each one of the following modes:**
B2_i	Passenger car (driver or passenger)
B2_ii	Pedestrian
B2_iii	Public transport (metro, bus, trolley bus, tramway)
	As a pedestrian, how much would you agree with each one of the following statements*:**

B3_i.	I walk for the pleasure of it
B3_ii	I walk because it is healthy
B3_iii	In short trips, I prefer to walk
B3_iv	I prefer taking public transportation (buses, metro, tramway, etc.) than my car
B3_iv	I walk because I have no other choice
C	**As a pedestrian, how much would you agree with each one of the following statements***:**
C1_i.	Crossing roads is difficult
C1_ii.	Crossing roads outside designated locations increases the risk of accident
C1_iii.	Crossing roads outside designated locations is wrong
C1_iv	Crossing roads outside designated locations saves time
C1_v	Crossing roads outside designated locations is acceptable because other people do it
C2_i	I prefer routes with singalised crosswalks
C2_ii	I try to make as few road crossings as possible
C2_iii	I try to take the most direct route to my destination
C2_iv	I prefer to cross diagonally
C2_v	I try to take the route with least traffic to my destination
C2_vi	I am willing to make a detour to find a protected crossing
C2_vii	I am willing to take any opportunity to cross
C2_viii	I am willing to make dangerous actions as a pedestrian to save time
D	**Compared to other pedestrians, how much do you agree that***:**
D_i	I am less likely to be involved in a road crash than other pedestrians
D_ii	I am faster than other pedestrians
D_iii	I am more careful than other pedestrians
E	**As a pedestrian, how often do you adopt each one of the following behaviours****:**
E1_i.	I cross diagonally
E1_ii	I cross at midblock at major urban arterials
E1_iii	I cross at midblock at urban roads
E1_iv	I cross at midblock in residential areas
E1_v	I cross at midblock when I am in a hurry
E1_vi	I cross at midblock when there is no oncoming traffic
E1_vii	I cross at midblock when I see other people do it
E1_viii	I cross at midblock when my company prompts me to do it
E1_ix	I prompt my company to cross at midblock
E1_x	I cross at midblock when there is a shop I like on the other side
E1_xi	I cross even though the pedestrian light is red
E1_xii	I walk on the pavement rather than on the sidewalk
E2_i	I cross between vehicles stopped on the roadway in traffic jams
E2_ii	I cross without paying attention to traffic
E2_iii	I am absent-minded while walking
E2_iv	I cross while talking on my cell phone or listing to music on my headphones
E2_v	I cross even though obstacles (parked vehicles, buildings, trees, etc.) obstruct visibility
E2_vi	I cross even though there are oncoming vehicles
F	**As a pedestrian, how much would you agree with each one of the following statements***:**
F1_i	Drivers are not respectful to pedestrians
F1_ii	Drivers drive too fast
F1_iii	Drivers are aggressive and careless
F1_iv	Drivers should always give way to pedestrians
F1_v	When there is an accident, it is the driver's fault most of the times
F1_vi	I let a car go by, even if I have right-of-way
* (1:never, 2: less than once a week, 3:once a week, 4: more than once a week, 5:every day)	

** (1:1-2 km, 2: 3-5 km, 3:5-20 km, 4: 20-50 km, 5: >50 km)

*** (1:strongly disagree, 2: disagree, 3:neither agree nor disagree, 4: agree, 5:strongly agree)

**** (1:never, 2: rarely, 3:sometimes, 4: often, 5:always)

3. Analysis methods

3.1. Principal Component Analysis

A Principal Component Analysis (PCA) is typically implemented on sets of selected variables or "indicators", in order to identify groups of variables ("components" or "dimensions"), to understand the structure of this set of variables and to reduce the dataset to a more manageable size and at the same time retain as much of the original information as possible.

A key issue is to determine the optimal number of components; the components that are retained were those that had an Eigenvalue above 1, indicating a large proportion of variance explained. Another criterion, however, concerns the interpretability of the components; it is preferable to maintain as few components as possible, explaining as much of the variance as possible. For example, it is possible that the Eigenvalue solution reveals several components, but the first few ones may already explain a sufficient share of the total variance in the data.

Standard principal components analysis assumes linear relationships between numeric variables. However, this assumption may not always stand, especially when dealing with discrete data. Another limitation of standard PCA is the adequate sample size requirement.

For these reasons, another approach is proposed, namely Categorical Principal Component Analysis (CATPCA), which fall within the broad family of optimal scaling techniques. With these techniques, discrete (nominal and ordinal) variables can be converted to "interval" variables, i.e. variables which are continuous within a given interval. As a result, nonlinear relationships between variables can be modelled (Muelman et al. 2004).

The first step of optimal scaling is the selection of the scaling and weighting level for the transformation of discrete variables into interval ones . Nominal, ordinal or spline weights can applied, in accordance to the nature of the examined variables, in order to preserve the type and order of the categories in the optimally scaled variable. Moreover, a 'grouping' or 'ranking' method can be applied for recoding the variables (Linting et al., 2007). The process results in the creation of new, transformed variables, which maintain the properties of the initial variables but are interval-continuous ones. Then, the CATPCA is applied as usual on the transformed (optimally scaled) variables / indicators.

3.2. Cluster Analysis

Cluster analysis is a similar technique to PCA, but the aim is to group individuals (respondents) instead of variables. Cluster analysis techniques are hierarchical or agglomerative. Hierarchical cluster analysis clusters (n) respondents on the basis of an algorithm starting from one single group to (n) groups. The observation of the statistical values of the algorithm for each level of clustering may lead to the identification of different clustering options; in this method, there is often no unique clustering solution, however useful insight is provided on the number and size of clusters.

K-means cluster analysis is an alternative method, in which respondents will be clustered in (k) predefined groups on the basis on their scores on the variables / indicators, in order to examine whether the indicators distinguish respondents in such groups. If this is the case, then the indicators may have an important effect as potential explanatory variables. If not, this would mean that all respondents tended to respond similarly to all questions.

In this research, instead of grouping respondents on the basis of their scores on the individual indicators, it will be attempted to group respondents on the basis of their scores on the components / dimensions estimated by the CATPCA.

4. Results

A descriptive analysis is the first step for the exploration of the human factors of pedestrian walking and crossing behaviour, as measured through the questionnaire. Subsequently, PCA and CATPCA techniques are implemented on the questionnaire responses, revealing the dimensions underlying the data. A cluster analysis of respondents on the basis on their scores on the dimensions identified, eventually allows the creation of pedestrian profiles.

4.1. Descriptive statistics

As shown in Figure 1, most participants have rather positive travel motivations (e.g. health and pleasure purposes have high positive scores). More than one third (36%) of the participants report that they walk because they do not have another choice.

B3	As a pedestrian, how much would you agree with each one of the following statements:	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
B3_i.	I walk for the pleasure of it	6%	4%	29%	43%	18%
B3_ii	I walk because it is healthy	3%	4%	18%	53%	22%
B3_iii	In short trips, I prefer to walk	1%	4%	10%	39%	46%
B3_iv	I prefer taking public transportation (buses, metro, tramway, etc.) than my car	15%	31%	26%	17%	11%
B3_iv	I walk because I have no other choice	19%	21%	24%	18%	18%

Fig.1. Distribution of travel motivations

Figure 2 summarises the responses on risk perceptions related to road crossing, value of time and opportunistic behaviour etc. Most pedestrians have positive attitudes and preferences (e.g. risk-conscious and compliant), as they tend to agree that crossing roads outside designated locations is risky and wrong, although they acknowledge that it saves time. Nevertheless, there is a non negligible proportion of pedestrians (around 15%) who disagree with these statements).

Moreover, the majority of pedestrians disagree with crossing diagonally and making dangerous actions to save time, while they agree with trying to minimize the number of crossings, taking the most direct route and taking the route with the least traffic. Again, there is a non-negligible proportion of pedestrians (20-35%) who are willing to take any crossing opportunity, make dangerous actions or prefer to cross diagonally.

C1	As a pedestrian, how much would you agree with each one of the following statements:	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
C1_i.	Crossing roads is difficult	10%	33%	25%	29%	3%
C1_ii.	Crossing roads outside designated locations increases the risk of accident	3%	13%	13%	53%	19%
C1_iii.	Crossing roads outside designated locations is wrong	7%	19%	22%	42%	10%
C1_iv	Crossing roads outside designated locations saves time	4%	15%	21%	38%	22%
C1_v	Crossing roads outside designated locations is acceptable because other people do it	19%	43%	24%	7%	7%

C2	As a pedestrian, how much would you agree with each one of the following statements:	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
C2_i	I prefer routes with singalised crosswalks	1%	8%	32%	50%	8%
C2_ii	I try to make as few road crossings as possible	3%	17%	26%	42%	13%
C2_iii	I try to take the most direct route to my destination	0%	1%	19%	44%	35%
C2_iv	I prefer to cross diagonally	6%	35%	29%	25%	6%
C2_v	I try to take the route with least traffic to my destination	1%	6%	25%	44%	24%
C2_vi	I am willing to make a detour to find a protected crossing	8%	24%	38%	25%	6%
C2_vii	I am willing to take any opportunity to cross	0%	22%	39%	28%	11%
C2_viii	I am willing to make dangerous actions as a pedestrian to save time	18%	38%	21%	21%	3%

Fig.2. Distribution of pedestrian perceptions, attitudes and preferences

In Figure 3, the sample statistics with respect to self-assessment and identity are presented. Very few pedestrians disagree with the self-assessment statements presented in the questionnaire. Interestingly, they all seem to agree or strongly agree that they are faster and safer than other pedestrians, as well as more careful than other pedestrians.

D	Compared to other pedestrians, how much do you agree that:	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
D_i	I am less likely to be involved in a road crash than other pedestrians	1%	10%	26%	50%	13%
D_ii	I am faster than other pedestrians	3%	14%	26%	38%	19%
D_iii	I am more careful than other pedestrians	1%	8%	26%	54%	10%

Fig.3. Distribution of pedestrian self-assessment and identity

Figure 4 summarises the participants' self-reported behaviour, compliance and risk-taking. There appear to be pedestrians with positive and compliant behaviour as well as pedestrians with less positive behaviour. Only 43% report that they never cross at mid-block in major urban arterials. On the other hand, most pedestrians will cross at mid-block when there is no oncoming traffic or between stopped vehicles in traffic congestion.

Even at different traffic conditions (e.g. urban roads vs. residential areas) there are different frequencies of different behaviours. There is lack of 'extreme' values in most questions; this may indicate that all pedestrians may adopt mid-block crossing under certain conditions. Nevertheless, no pedestrian responded crossing without paying any attention to traffic. It may be interesting to note, that pedestrians report not to be influenced by their company as regards mid-block crossing.

E1	As a pedestrian, how often do you adopt each one of the following behaviours:	Never	Rarely	Sometimes	Often	Always
E1_i.	I cross diagonally	10%	33%	33%	21%	3%
E1_ii	I cross at midblock at major urban arterials	43%	29%	18%	10%	0%
E1_iii	I cross at midblock at urban roads	11%	28%	36%	22%	3%
E1_iv	I cross at midblock in residential areas	1%	11%	24%	47%	17%
E1_v	I cross at midblock when I am in a hurry	3%	11%	25%	46%	15%
E1_vi	I cross at midblock when there is no oncoming traffic	3%	7%	17%	39%	35%
E1_vii	I cross at midblock when I see other people do it	21%	33%	35%	10%	1%
E1_viii	I cross at midblock when my company prompts me to do it	20%	28%	24%	23%	6%
E1_ix	I prompt my company to cross at midblock	21%	37%	24%	14%	4%
E1_x	I cross at midblock when there is a shop I like on the other side	11%	37%	24%	15%	13%
E1_xi	I cross even though the pedestrian light is red	13%	28%	31%	28%	1%
E1_xii	I walk on the pavement rather than on the sidewalk	4%	19%	32%	29%	15%

E2	As a pedestrian, how often do you adopt each one of the following behaviours:	Never	Rarely	Sometimes	Often	Always
E2_i	I cross between vehicles stopped on the roadway in traffic jams	1%	15%	25%	47%	11%
E2_ii	I cross without paying attention to traffic	63%	31%	7%	0%	0%
E2_iii	I am absent-minded while walking	21%	39%	28%	8%	4%
E2_iv	I cross while talking on my cell phone or listing to music on my headphones	7%	18%	42%	26%	7%
E2_v	I cross even though obstacles (parked vehicles, buildings, trees, etc.) obstruct visibility	21%	26%	32%	18%	3%
E2_vi	I cross even though there are oncoming vehicles	19%	42%	29%	8%	1%

Fig.4. Distribution of pedestrian behaviour, compliance and risk-taking

Finally, in Figure 5, the participants' responses with respect to their opinion on drivers are summarised. It is observed that pedestrians tend to agree with negative opinions on drivers. However, they do not believe that accidents are mostly drivers' fault.

F1	As a pedestrian, how much would you agree with each one of the following statements:	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
F1_i	Drivers are not respectful to pedestrians	0%	13%	19%	49%	19%
F1_ii	Drivers drive too fast	1%	7%	32%	47%	13%
F1_iii	Drivers are aggressive and careless	1%	6%	26%	49%	18%
F1_iv	Drivers should always give way to pedestrians	3%	14%	24%	36%	24%
F1_v	When there is an accident, it is the driver's fault most of the times	1%	31%	44%	21%	3%
F1_vi	I let a car go by, even if I have right-of-way	7%	26%	36%	29%	1%

Fig.5. Distribution of pedestrians' opinion on drivers

4.2. Dimensions of pedestrian human factors

A PCA was initially implemented on the 52 questions of the survey. The initial solution suggested that there are 14 components explaining in total 74% of the total variance. It was noted, however, that the 5 first components already explained almost 50% of the total variance; moreover, beyond component 9 there were no component 'loadings' higher than 0.5, revealing that the remaining components are formed on the basis of relatively weak relationships between variables. It was also revealed that, as usual, the first component was 'overloaded' with several indicators.

In order to obtain better insight on the dimensionality of the responses, the 'large' questionnaire sections C and E were examined separately, and it was indicated that each one of them is described by no more than one or two components. Moreover, an outlier analysis was performed, revealing a few specific individuals who systematically deviated from the overall trends in several questions from different sections - to the extent that separate components were formed only for these few individuals - and therefore it was decided to eliminate these individuals from the analysis. It is noted that the mean component scores of the analysis were then assigned to these respondents, so that they would not be excluded from the subsequent analyses.

Subsequently, CATPCA was implemented on the final sample, by testing three scenarios: two components, three components and five components. The testing of two components revealed a significant amount of variance explained, however with strong loadings of indicators C and E within component 1, with only a few loadings exceeding 0.4 on component 2 (again mostly from parts C and E). It was therefore confirmed that there is one major component on parts C and E of the questionnaire.

The testing of three components revealed a significant additional amount of variance explained, as the Eigenvalue of the third component was also high. However, the component loadings suggested that the introduction of a third component, with an indicator from part B of the questionnaire emerging as a predictor for the first time.

The testing of five components revealed an even more balanced identification of components. First, it was noted that all five components have high Eigenvalues, and therefore contribute significantly to the total variance explained. Moreover, there was a clearer 'contrast' between the two first components, including indicators from sections C and E, and the third component was strongly based on part B of the questionnaire. The remaining two components (4 and 5) were weak and not directly interpretable, however. For a detailed description of these analysis steps the reader is referred to Papadimitriou et al. (2015a).

The optimal solution selected was one with the 3 first components of the five-component analysis, explaining 65% of the total variance. Table 3 shows these 3 components / dimensions and the related indicators with loadings higher than 0.40. These three dimensions can be summarized as follows:

- Dimension 1, "risk taking and optimization": this component brings together elements of the questionnaire related to risk-taking behavior, namely the tendency to cross at mid-block, diagonally, at the presence of oncoming vehicles, etc., and also related to optimization of the trip, namely the tendency to minimize crossings, save time, avoid detours etc. These responses also appear to be correlated with low risk perception (e.g. negative scores for "crossing outside designated locations is difficult", or "it increases the risk of accident").
- Dimension 2, "conservative and public transport user": this component is rather opposed to the optimization patterns identified in component 1, as it brings together the tendency not to minimize crossings and not cross diagonally (e.g. not avoiding detours or delays), and is also correlated with increased perceived difficulty of road crossing. These responses are correlated with high and frequent pedestrian activity, but most importantly with frequent use and preference to use public transportation.
- Dimension 3, "pedestrian for pleasure", also reflects increased pedestrian activity, similar to that of component 2, but has distinctive high scores in "walking for pleasure" and "walking for health", "crossing at mid-block to see a shop" etc. Finally, it is correlated with increased perception of drivers being at fault in vehicle-pedestrian accidents.

Table 3. Dimensions of pedestrian behaviour: CATPCA component loadings and eigenvalues

Dimension 1: Risk taker & optimizer (Eigenvalue=9.06)	
Crossing roads outside designated locations increases the risk of accident	-.568
Crossing roads outside designated locations is wrong	-.509
Crossing roads outside designated locations is acceptable because other people do it	.418
I prefer to cross diagonally	.633
I am willing to make a detour to find a protected crossing	-.564
I am willing to take any opportunity to cross	.636
I am willing to make dangerous actions as a pedestrian to save time	.526
I am faster than other pedestrians	.473
I cross diagonally	.674
I cross at midblock at major urban arterials	.579
I cross at midblock at urban roads	.739
I cross at midblock in residential areas	.723
I cross at midblock when I am in a hurry	.825
I cross at midblock when there is no oncoming traffic	.602
I cross at midblock when I see other people do it	.467
I cross at midblock when my company prompts me to do it	.575
I prompt my company to cross at midblock	.746
I cross even though the pedestrian light is red	.593
I cross between vehicles stopped on the roadway in traffic jams	.658
I cross even though obstacles (parked vehicles, buildings, trees, etc.) obstruct visibility	.548
I cross even though there are oncoming vehicles	.683
Dimension 2: Conservative & public transport user (Eigenvalue=4.40)	

Weekly travel by Public transport	.698
Weekly travel by Pedestrian	.470
Weekly travel by Passenger car	-.534
Weekly Km of travel by Passenger car	-.475
Weekly Km of travel by Public transport	.724
I prefer taking public transportation than my car	.493
Crossing roads is difficult	.558
I try to make as few road crossings as possible	-.463
I prefer to cross diagonally	-.503
I am less likely to be involved in a road crash than other pedestrians	-.452
Dimension 3: Pedestrian for pleasure (Eigenvalue=3.67)	
Weekly travel by Pedestrian	.570
Weekly travel by Passenger car (driver or passenger)	-.593
WeeklyKm of travel by Passenger car (driver or passenger)	-.534
WeeklyKm of travel by Pedestrian	.583
I walk for the pleasure of it	.562
I walk because it is healthy	.628
I prefer routes with singalised crosswalks	.419
I am willing to make a detour to find a protected crossing	.417
I cross at midblock when there is a shop I like on the other side	.425
When there is an accident, it is the driver’s fault most of the times	.478

4.3. Pedestrian ‘profiles’ (clusters)

As a final step of the analysis, pedestrian ‘profile’ identification was attempted. First, a hierarchical cluster analysis was implemented on the whole sample. The dendrogram of the hierarchical cluster analysis of Figure 6 provides a final confirmation towards the optimisation of group identification. Two main groups / branches are clearly visible, but some ‘noise’ is also visible at the bottom right, where there are three ‘marginal’ branches including very few individuals each, namely the respondents who were identified as ‘outliers’ also in the PCA. It was confirmed that these individuals should be removed for a more meaningful solution.

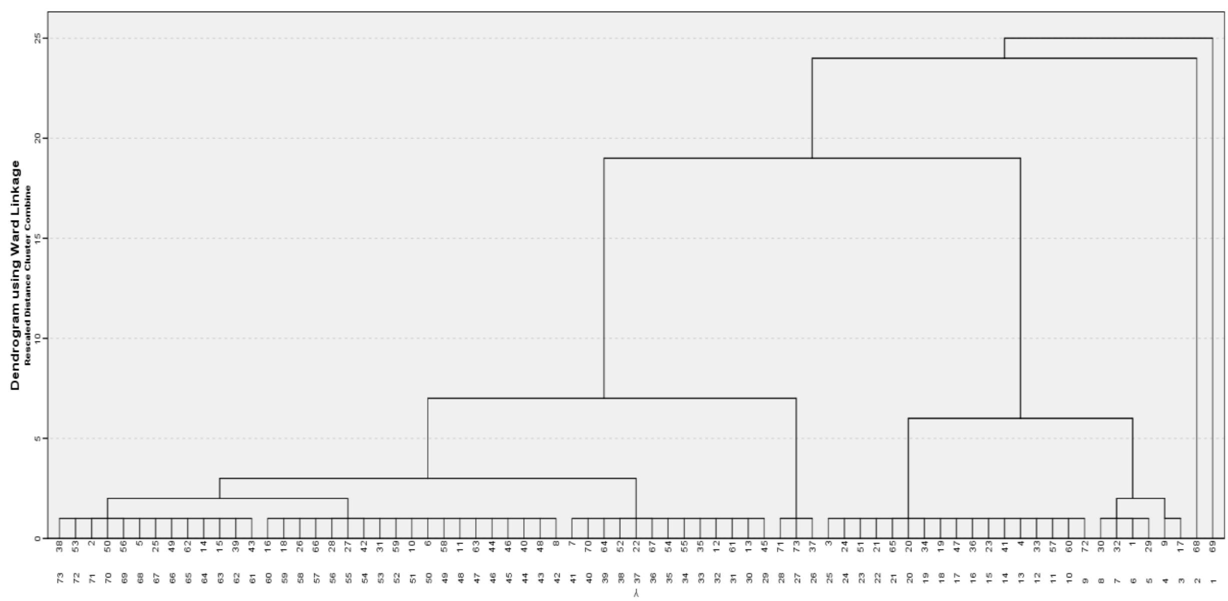


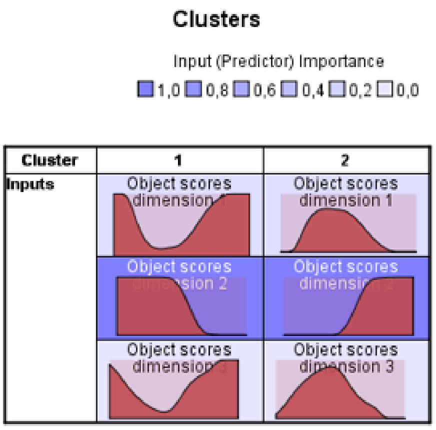
Fig.6. Hierarchical cluster analysis dendrogram of questionnaire responses

For clustering the respondents, the scores of the three first components (dimensions) of the 5-component solution were used. Figure 7 shows that two clusters of respondents are formed, and in this case they both include an adequate sample of participants (distribution of almost 40%-60%). The complementarity between clusters on the basis of the component scores is highlighted, with the 2nd component appearing to be a strong predictor of cluster membership.

The cluster centers (defined as the mean score within each cluster for each component) are also presented in Figure 7. It is found that cluster 1 includes individuals with positive scores on components 1 and 2, and negative scores on component 3, while cluster 2 includes individuals with the opposite sign and magnitude. Attempting a first meaningful interpretation of the groups, group 1 seems to include high responses on parts C and E, which corresponds to high risk taking attitudes and behaviours, and low responses on part B, which corresponds to low travel motivations with respect to walking frequency, health and pleasure purposes etc.

It is noted that the positive scores of cluster 1 on dimension 2 indicate more negative attitudes and behaviours, as it is reminded that dimension 2 was formed on negative loadings on the related indicators, i.e. negative scores on these indicators (e.g. "disagree). Consequently, the two groups of pedestrians can be described as follows:

- Group 1 "Negative attitudes and behavior / weak walking motivations": pedestrians in this group have rather negative attitudes, perceptions and behavior, as they have high scores on risk-taking and trip optimization (i.e. low delay and detour acceptance) and low scores on illegal crossing perceived as dangerous or wrong. Moreover, they rarely use public transport and do not walk for pleasure or health purposes, which indicates weak motivations for walking.
- Group 2 "Positive attitudes and behavior / strong walking motivations": pedestrians in this group have positive attitudes, perceptions and behavior, as they have low scores on risk-taking behaviour and non-compliance and high scores on risk perception. Moreover, they frequently use public transport and walk for pleasure or health purposes, which indicates strong motivations for walking.



Final Cluster Centers		
	Cluster	
	1: Negative attitudes and behaviours and weak motivations	2: Positive attitudes and behaviours and strong motivations
Object scores dimension 1	,23	-1,11
Object scores dimension 2	,16	-,77
Object scores dimension 3	-,22	1,09

Fig.7. Two-step cluster analysis of questionnaire responses on three-dimensions scores - Final cluster centres

5. Conclusions

This research concerns an in-depth analysis of human factors of pedestrian walking and crossing behaviour on the basis of specific research hypotheses derived from the literature, and a specially designed questionnaire. The descriptive and exploratory analysis of the questionnaire responses clearly suggested that there are few patterns of pedestrians' attitudes, perceptions and behaviours. In all sections of the questionnaire and in all indicators, there were both positive and negative responses.

From the preliminary analysis, it was suggested that the identification of the dimensionality of the questionnaire responses and the estimation of human factors of pedestrian crossing behaviour as "components" is not straightforward. Several steps were taken in order to identify the optimal solution. Overall, there is strong indication that there are no more than 3 meaningful dimensions / components in the data, the two first strongly related to pedestrian risk-taking attitudes and perceptions, and the third one strongly related to pedestrians walking motivations. There is also strong indication that specific respondents are in fact 'outliers' in the analysis, creating 'artificial' components and groups which are beyond the core components and groups in the data.

The clustering of pedestrians into meaningful groups sharing common scores in underlying dimensions was more straightforward, with two 'opposite' groups easily emerging: on the one hand "positive" (i.e. non risk taking, compliant, motivated) pedestrians, and on the other hand "negative" (risk-taking, impatient, unmotivated) pedestrians.

One key finding is that, the research hypotheses on the human factors of pedestrian crossing behaviour, as reflected in the structure of the questionnaire, are not confirmed by the above results. In particular, it was assumed that there were five factors of pedestrian behaviour, each one corresponding to one section of the survey questionnaire (B, C, D, E, and F). The descriptive analysis of the data also revealed diversity in the responses. However, the statistical analysis did not confirm these indications, suggesting that the underlying human factors dimensions are in fact few, and the 'profiles' or types of pedestrians even fewer.

These findings should be considered in light of the limitations of the present research. The survey sample was rather small, as the questionnaire respondents were participants of a larger observational study on pedestrian behaviour (Papadimitriou et al., 2015a) and therefore recruitment was not as easy as in a typical questionnaire survey. For the same reasons, not all age groups are adequately represented in the sample; older pedestrians are not included, and middle-aged pedestrians are under-represented. Extending this questionnaire survey to a larger and more representative sample may reveal additional dimensions and aspects of pedestrian behaviour, and additional 'profiles' of pedestrians.

The present research confirms that human factors are potentially important determinants of pedestrian behaviour, as the observed diversity in attitudes, perceptions and declared behaviours, and the strong "contrast" between the pedestrian profiles identified would certainly have an effect on pedestrian walking and crossing behaviour. The association of human factors identified with observed pedestrian behaviour models has been attempted in a few recent studies with this dataset (Papadimitriou et al., 2015b) or with other datasets (e.g. Granié, 2007; Sisiopiku & Akin, 2010), with promising results on the explanatory potential of human factors on pedestrian behaviour.

Acknowledgements

This paper is based on a research project implemented within the framework of the Action «Supporting Postdoctoral Researchers» of the Operational Program "Education and Lifelong Learning" (Action's Beneficiary: General Secretariat for Research and Technology), and is co-financed by the European Social Fund (ESF) and the Greek State. The authors would like to thank Dr. Marie-Axelle Granié from IFSTTAR for her valuable comments and suggestions in the earlier stages of this research.

References

Bernhoft I.M., Carstensen G. (2008). Preferences and behaviour of pedestrians and cyclists by age and gender. Transportation Research Part F 11, pp. 83-95.

Diaz, E.M. (2002). Theory of planned behaviour and pedestrians' intentions to violate traffic regulations. Transportation Research Part F 5, 169–175.

Evans D., Norman P. (1998). Understanding pedestrians' road crossing decisions: an application of the theory of planned behaviour. Health Education Research 13 (4), pp. 481-489.

Fuller R., Santos J.A. (2002). Human Factors for Highway Engineers. Elsevier Science, UK.

Granié M.A. (2007) Gender differences in preschool children's declared and behavioural compliance with pedestrian rules. Transportation Research Part F: Traffic Psychology and Behaviour 10 (5), pp. 371-381.

Grannie M.A., Pannetier M., Guého L. (2013). Developing a self-reporting method to measure pedestrian behaviours at all ages. Accident Analysis and Prevention 50 (2013) 830– 839.

Hine J. (1996). Pedestrian travel experiences: Assessing the impact of traffic on behaviour and perceptions of safety using an in-depth interview technique. Journal of Transport Geography 4 (3), 179-199.

Holland, C.A., Hill, R., 2010. Gender differences in factors predicting unsafe crossing decisions in adult pedestrians across the lifespan: a simulation study. Accident Analysis and Prevention 39(2), pp. 224-237.

Linting, M., Meulman, J.J., Groenen, P.J.F., van der Koojj, A.J. (2007). Nonlinear principal components analysis: Introduction and application. Psychological Methods, Vol 12(3), Sep 2007, 336-358.

Meulman J.J, van Der Kooij A.J., HeiserKaplan W.J. (2004). Principal Components Analysis with Nonlinear Optimal Scaling Transformations for Ordinal and Nominal Data. In: Kaplan D. (ed): The SAGE Handbook of Quantitative Methodology for the Social Sciences. SAGE Publications, USA.

Oxley, J., Fildes, B., Ihsen, E., Charlton, J., and Days, R. (1997). Differences in traffic judgements between young and old adult pedestrians. Accid. Anal. and Prev. 29 (6), pp. 839-847.

Papadimitriou E., Yannis G., Golias J. (2009). A critical assessment of pedestrian behaviour models. Transportation Research Part F 12 (3), pp. 242-255.

Papadimitriou E. (2012). Theory and models of pedestrian crossing behaviour along urban trips. Transportation Research Part F 15 (1), pp. 75-94.

Papadimitriou E., Theofilatos A., Yannis G. (2013). Patterns of pedestrian attitudes, perceptions and behaviour in Europe. Safety Science 53, pp. 114-122.

Papadimitriou E., Yannis G., Lassarre S. (2015a). Final report on models for pedestrian behaviour and safety. Deliverable 7.1 of the PEDMOD project of the Action "Supporting Post-Doctoral Researchers" of the Operational Program "Education and Lifelong Learning", General Secretariat for Research and Technology, Greece.

Papadimitriou E., Lassarre S., Yannis G., (2015b). Introducing human factors in pedestrian crossing behaviour models. Proceedings of the 94th Annual meeting of the Transportation Research Board, Washington DC, January 2015.

Sisiopiku, V.P., Akin, D., 2003. Pedestrian behaviours at and perceptions towards various pedestrian facilities: an examination based on observation and survey data. Transportation Research Part F: Traffic Psychology and Behaviour 6, 249– 274.

Yagil, D., 2000. Beliefs, motives and situational factors related to pedestrians' self reported behaviour at signal-controlled crossings. Transportation Research Part F: Traffic Psychology and Behaviour 3 (1), 1–13.