Introducing human factors in pedestrian crossing behaviour models

Eleonora Papadimitriou[a1], Sylvain Lassarre[b], George Yannis[a]

[a] National Technical University of Athens
Department of Transportation Planning and Engineering

[b] IFSTTAR - Institut français des sciences et technologies des transports, de l'aménagement et des réseaux

Abstract

The objective of this research is the development of pedestrian crossing choice models on the basis of road, traffic and human factors. For that purpose, a field survey was carried out, in which a panel of 75 pedestrians were asked to take 8 short walking trips (each one corresponding to a different walking and crossing scenario) in the Athens city centre in Greece, allowing to record their crossing behavior in different road and traffic conditions. The same individuals were asked to fill in a questionnaire on their travel motivations, their mobility characteristics, their risk perceptions and preferences with respect to walking and road crossing, their opinion on drivers etc. The walking and crossing scenarios‟ data were used to develop mixed sequential logit models of pedestrian behavior on the basis of road and traffic characteristics. The modeling results showed that pedestrian crossing choices are significantly affected by road type, traffic flow and traffic control. The questionnaire data were used to estimate human factors (components) of pedestrian crossing behavior by means of principal component analysis. The results showed that three components of pedestrian crossing behavior emerge, namely a "risk-taking and optimization" component reflecting the tendency to cross at mid-block in order to save time etc., a "conservative" component, concerning individuals with increased perceived risk of mid-block crossing, who also appear to be frequent public transport users, and a "pedestrian for pleasure" component, bringing together frequent pedestrians, walking for health or pleasure etc. The introduction of these components as explanatory variables into the choice models resulted in improvement of the modeling results, indicating that human factors have additional explanatory power over road and traffic factors of pedestrian behavior. Therefore, the development of integrated choice and latent variables models appears to be an appropriate field for further research.

Key-words: pedestrian behavior; human factors; mixed logit models.

1. Background and objectives

The analysis of pedestrians crossing behaviour in urban areas may assist in the understanding of the way pedestrians interact with the road and traffic environment, as well as with other pedestrians, and the way they balance the need for comfort and safety with the cost of delays, within the framework of existing traffic rules (Das et al., 2005). Eventually, it may assist in the better adjustment of urban road networks to pedestrians' needs and the more accurate estimation of pedestrians road accident risk exposure in urban areas (Lassarre et al., 2007) and thus to the improvement of pedestrians safety.

Although signalized junctions provide pedestrians a protected crossing phase, most pedestrians tend to prefer using the available traffic gaps for crossing (Hamed, 2001). Moreover, mid-block crossing and diagonal crossing are common practice among pedestrians aiming to save travel time or distance (Chu et al., 2003). Because of their flexibility and adaptability, pedestrians generally experience smaller delays compared to other road users, but increased road accident risk exposure (Grayson, 1987).

Existing research on pedestrians crossing behaviour in urban areas is extensive and largely concerns gap acceptance models, in which each pedestrian is associated with a critical gap for road crossing (Himanen & Kulmala, 1988; Sun et al., 2003; Oxley et al., 2005). In several researches, a level of service approach is implemented for road crossing, in which the difficulty to cross is used as a measure of effectiveness for pedestrian level of service (Sarkar, 1995; Balted & Chu, 2002). Moreover, pedestrians' crossing choices among a set of discrete alternatives are often modelled on the basis of utility theory (Chu et al., 2003; Muraleetharan et al., 2004; Papadimitriou, 2012).

A distinct part of existing research on pedestrian crossing behaviour is devoted to analyses of psychological, attitudinal, perceptual and motivational factors (Evans & Norman, 1998; Diaz, 2002; Bernhoft & Carstensen, 2008). However, human factors are seldom incorporated in pedestrian behaviour and safety models, so that the explanatory power of these factors can be tested. It is common to analyse the observed behaviour of pedestrians in relation to road and traffic characteristics, or the self-reported behaviour, attitudes and perceptions on the basis of questionnaire surveys, but the entire set of potential determinants has not been jointly explored in the existing studies.

The objective of this paper is the analysis of pedestrian crossing behaviour along entire trips in urban road networks, with particular emphasis on the introduction of human factors (pedestrians attitudes, perceptions etc.) in the potential determinants. More specifically, this research aims to develop choice models for estimating the probability to cross at each location along a pedestrian trip in relation to roadway design, traffic flow and traffic control, as well as human factors.

This paper starts with the description of the data collection scheme used for the purposes of this research, namely a combination of field observations and questionnaire survey. Subsequently, the analysis techniques are presented, which include mixed sequential logit models for pedestrian crossing behavior (applied on the field observations data) and categorical principal components analysis (applied on the questionnaire data). The results section follows, including the sequential choice models of pedestrian crossing behavior, the components of pedestrian human factors, and the introduction of these components into the choice models. The paper ends with a discussion of the findings, also in light of the next steps of the research.

2. Data collection

In this research, a particular data collection scheme was implemented. Pedestrians were followed along urban trips, and their crossing behaviour was recorded, together with features of the road environment and the traffic conditions. Furthermore, they were asked to fill in a questionnaire on their attitudes, perceptions and behaviour as regards road crossing and accident risk. The design of both the field survey and the questionnaire were based on an exhaustive review of the literature, leading to the formulation of specific research hypotheses to be tested.

2.1. Survey design

From the review of the literature it may be concluded that several road and traffic factors affect pedestrians road crossing choices, although research findings are not always consistent. The research hypotheses of the field survey design are as follows:

- Road type: in residential zones (minor urban roads), pedestrians will choose the shortest path (e.g. cross diagonally), due to the lack of constraints and vehicle-pedestrian interaction. On major urban arterials, on the other hand, the constraints (traffic, speed of traffic, number of lanes) are such that all pedestrians will opt for a protected crossing at signalized junction. In a mixed urban area (e.g. secondary urban roads), more variation is expected in pedestrian crossing behaviour.
- Traffic flow: when there is no traffic, pedestrians will choose the shortest path (e.g. cross diagonally), due to the lack of constraints and vehicle-pedestrian interaction. At low traffic there is increased probability of crossing at mid-block or diagonally, while at high traffic there is increased probability of seeking for a protected crossing at junction. Finally, at congestion pedestrians are also likely to cross at midblock or diagonally, „in between" stopped vehicles.
- Traffic control: the presence of traffic signal control leads to increased probability of crossing at junction.
- Infrastructure design, obstacles and barriers: obstacles, such as illegally parked vehicles on the sidewalk, roadside barriers and guardrails, or specific local infrastructure design elements (e.g. a „gap" in roadside guardrails, a painted crosswalk at midblock, a change in sidewalk width) may lead pedestrians to a deterministic choice.

Given that there are correlations between several of the above road and traffic factors considered in the present research, not all combinations of these factors are meaningful. This naturally leads to a fractional factorial design of the field survey, intuitively produced by eliminating all the combinations that are not applicable in urban transportation networks. For example, uncontrolled crosswalks are not applicable in urban arterials, signal-controlled crosswalks are not applicable in residential roads, high traffic volumes and congestion conditions are not encountered in residential areas etc. On the basis of the above, the field survey design consists of three walking conditions and eight crossing scenarios:

- Crossing an main urban road with signal controlled and uncontrolled crosswalks: scenarios (i) and (viii);
- Crossing a minor (residential) road with or without marked crosswalks: scenarios (ii), (v), (vi) and (vii);
- Crossing a major urban arterial with signal controlled crosswalks: scenarios (iii) and (iv).

The selected field survey site is located at the Athens central area, from the Evangelismos metro station to the Kolonaki square. The area (see Figure 1) includes all three road types examined in the present research: a major urban arterial (Vas.Sofias Ave.), and its parallel two way main urban road (Patriarchou Ioakeim str.), with numerous commercial and recreation uses. These two roads are separated by a grid network of minor / residential roads, with mainly offices and residential uses. The major urban arterial and the main urban road may demonstrate all types of traffic conditions during the day, from free flow to congestion, while the minor / residential roads have mainly low to moderate traffic throughout the day.

Eight survey scenarios were developed, covering the options identified in the fractional factorial design. The eight survey scenarios were developed so that the choice sets for crossings can be clearly defined, as shown in Figure 1.

Figure 1 to be inserted here

2.2. Questionnaire design

For the development of the questionnaire, several questionnaires from the existing literature were studied. 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 research hypotheses can be summarized as follows:

- Demographics: Younger and male pedestrians are more risk-taking and less compliant to traffic rules related to road crossing. Low income, perceived social inequality and the lack of alternatives to walking may lead pedestrians to more aggressive and less compliant behaviour.
- Travel motivations: a positive relationship between walking frequency / distance travelled and crossing behaviour is assumed. Moreover, pedestrians walking for health / recreation purposes are likely to be less risk-taking and more safety conscious.
- Risk perception and value of time: there may be different types of pedestrians; ones that minimize the number of crossings and increases the length of the path

in order to avoid vehicle / pedestrian interaction, and others who maximize the number of crossings in order to reduce the length of the pedestrian path;

- Traffic behaviour and 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 (i.e. some pedestrians may „follow" the crossing choices of others, while others may prompt their company to a specific behaviour, opinion towards drivers (i.e. pedestrians with negative opinion on drivers are more likely to be careful and compliant).

On the basis of the above research hypotheses, the questionnaire includes 6 sections, as shown in Table 1:

- Section A: Demographics
- Section B: Mobility and travel motivations
- Section C: Attitudes, perceptions and preferences
- Section D: Self-assessment and identity
- Section E: Behaviour, compliance and risk taking
- Section F: Opinion on drivers

Table 1 to be inserted here

2.3. Survey procedure

A pilot data collection took place on July 2013, including 7 participants. The first wave of data collection took place in the period September – October 2013 concerning 30 more participants. The second wave of data collection tool place in December 2013 concerning 37 more participants. Half of the participants first filled in the questionnaire and then carried out the field experiment, and half of the participants fist carried out the field experiment and then filled in the questionnaire.

Given that the walking scenarios are fixed and similar for all the participants, there are two types of data recorded in the field survey:

- Static data: these concern the characteristics of the trips in terms of street names, road geometry and traffic control available in each case. These only need to be recorded once and are the same for all participants;
- Dynamic data: these concern the walking and crossing characteristics of the participants, they were recorded in real time conditions while following the pedestrian, and can be further distinguished into:
- – Data recorded for each road link, e.g. walking time and length, traffic volume, number and duration of crossing attempts etc.
- – Additional data recorded for road links with a crossing, e.g. crossing location (junction or mid-block), crossing type (diagonal), signal display (red / green), etc.

3. Analysis methods

3.1. Parameterisation of pedestrian trips and crossings

First, a parameterisation of pedestrian crossing behaviour is used, on the basis of existing work (12). This is based on a topological approach of urban road networks and pedestrian trips, on the basis of which the number and type of crossings along a pedestrian trip, as well as the set of choice alternatives for each road crossing can be determined.

In particular, the survey scenarios were designed so that only one crossing of interest will take place for each scenario, namely a „primary‟ crossing. Primary crossings have been defined in previous research (Lassarre et al., 2007; Papadimitriou, 2012) as crossings that take place across the pedestrian trajectory and their choice is stochastic (i.e. pedestrian may choose from a number of alternative locations along the trajectory for changing side of the road). On the other hand, there are crossings whose choice is deterministic, referred to as „secondary‟ crossings, as they take place „along‟ the trip trajectory, without changing side of the trajectory. A more detailed discussion of primary and secondary crossings is beyond the scope of this paper, as they are discussed in detail with many examples in several recent papers (Lassarre et al., 2007; Papadimitriou, 2012; Papadimitriou et al. 2012). In Figure 1, the red arrows represent examples of the „primary‟ crossings to be examined.

3.2. Mixed sequential choice models

A probabilistic discrete choice is involved in determining the location of each primary crossing from the alternatives of the examined scenario (choice set). A utility function is associated with each crossing alternative (i) for a pedestrian (n), as follows (Ben-Akiva & Lerman, 1986):

$$U_{in} = V_{in} + \varepsilon_{in}$$

Where $V_{in} = \beta' X_{in}$ is the systematic (deterministic) part of the utility, and ε_{in} is the random part of the utility function.

According to random utility theory, choice behaviour is based on the assessment of the known alternatives of the choice set. However, pedestrian crossing behaviour along a trip is a dynamic process, in which the alternatives of the choice set for each primary crossing become gradually available while the pedestrian moves along the road links of the choice set. Consequently, it can be assumed that, on each road link of the choice set, the pedestrian will assess the available crossing alternatives and decide whether to cross or not. If no crossing takes place on this link, the alternatives of the next road link are assessed and so on, until a crossing alternative is chosen, and the rest of the choice set does not need to be considered (Papadimitriou, 2012).

This process corresponds to a sequential choice behaviour, in which no prior knowledge of the road network or the trip conditions is assumed. Three alternatives are thus considered for each road link of each scenario, namely a 'crossing at junction' option (J), a 'crossing at mid-block' option (MB) and a 'no crossing' option (No). The crossing choice is therefore modelled by means of sequential multinomial logit models, as shown in Figure 2.

However, sequential choices of a group of individuals (panel data) can not be considered as independent. In fact, two types of dependence may be involved. The first one concerns individual-specific heterogeneity (Wooldridge, 2005), i.e. random variation resulting from the fact that these choices are repeated observations by the different individuals. The second one concerns state dependence (Honoré & Kyriazidou, 2000), due to the fact that each choice is made on different „states" (i.e. situations) of the same process, and thus the choice in the previous state may affect the choice of the current state (i.e. a typical 1st order Markov process).

Therefore, two types of extensions may need to be incorporated in the utility function of the model, which is now considered to apply to each „state" T=t (a separate „state" corresponding to each link of the choice set) as follows:

$U_{int} = \beta_i' X_{int} + \gamma y_{n,t-1} + \alpha_n + \varepsilon_{int}$

Where $y_{n,t-1}$ is the choice made in the previous „state" T= t-1 (state dependence), α_n is unobserved heterogeneity, which may be fixed or random, and ε_{int} the random part of the utility. The consideration of random heterogeneity results in a mixed logit model.

Figure 2 to be inserted here

3.3. Categorical Principal Component Analysis

Human factors of pedestrian behaviour are typically examined by means of questionnaire surveys, in which the responses to a set of actual questions (observed variables or „indicators") are used to estimate an unobserved or "latent" variable (often referred to as „factor" or „component"), such as risk perception, risk proneness, travel motivation, attitude towards walking etc. The latent variable is based on a linear or (more seldom) non-linear combination of the observed variables, resulting in a "score" reflecting the latent variable.

One family of techniques aiming to estimate latent variables are „component" analysis techniques, which seek to reveal underlying „components" (or „factors") structured on the basis of a relatively thematically organised set of indicators.
Standard principal components analysis assumes linear relationships between numeric variables. However, this assumption may not always stand, especially when dealing with discrete data. Categorical Principal Component Analysis (CATPCA), which falls within the broad family of optimal scaling techniques, converts discrete (nominal and ordinal) variables to "interval" variables, i.e. variables which are continuous within a given interval. The optimal-scaling approach allows variables to be scaled at different levels, and categorical variables are optimally quantified in a specified dimensionality. As a result, nonlinear relationships between variables can be modelled (Meulman 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 on the transformed (optimally

scaled) variables / indicators, in order to reduce the dimensionality of the dataset to a predefined number of dimensions or components.

4. Results

4.1. Estimation of human factors

In this research, optimal scaling was applied on the 51 variables of the questionnaire, which were defined as multiple ordinal, resulting in a „ranking" discretization method. By discretization, the discrete ordered values are transformed to interval-continuous on the basis of a ranking of their values. Components from the 51 optimally scaled variables are then extracted on the basis of the eigenvalue>1 criterion. This results in 3 components, explaining 65% of the total variance, as shown in Table 2. It is noted that component loadings lower than 0.40 have been deleted, in order to make the components interpretation more straightforward.

The three components can be summarized and tentatively labeled as follows:

- Component 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").
- Component 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.
- Component 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.

As a next step, it was examined whether the component scores for each pedestrian are significant explanatory variables of pedestrian crossing choices.

Table 2 to be inserted here

4.2. Development of a choice model with road, traffic and human factors

A mixed sequential logit model was fitted to the survey data. For each road link of each walking scenario, three options are available: "cross at mid-block", "cross at junction" and "not cross at all". In this type of choice modelling, the utility of each alternative is considered conditional on the availability of the alternative, i.e. it is explicitly indicated which alternatives are available in each case. For example, on the first road link of a trip, all three alternatives are available. Moreover, for each scenario, if the pedestrian did cross on one link, then the crossing alternatives are not available in the remaining links of the scenario (as each scenario corresponds to only one primary crossing). Similarly, if the pedestrian is on the last link of the scenario and has not already crossed, then the „no crossing‟ option is not available.

The modelling of the field survey data was carried out as follows: First, an „empty‟ model was fitted, including only the alternative-specific constants (ASC). Then, the best-fitting constrained model was found, including the statistically significant explanatory road and traffic variables. Variables can be: (i) generic, i.e. with a common B coefficient for all alternatives, or (ii) alternative-specific, i.e. with different B coefficients for each alternative. Typically, characteristics of the choice-maker are introduced as generic variables (e.g. gender, age), while characteristics („attributes‟) of the alternatives are introduced as alternative-specific variables (e.g. traffic flow etc.). In addition to these fixed effects, a random „panel‟ effect was examined, in order to capture heterogeneity due to unobserved differences between respondents.

The models were fit by using the Biogeme software for choice modelling (Bierlaire, 2003). The best-fitting model (not presented here) was one with alternative-specific parameters for road type, traffic flow, traffic signal and barriers. It also includes a „state-dependence‟ effect, reflected by the first road link, the skipping of one crossing opportunity etc. The significance level considered acceptable is 90%. A likelihood ratio test leads to accept the model.

More specifically:

- „State dependence‟ (B0_first, B1_first, B0_skip1): the first road link is more likely to be chosen for crossing compared to not crossing. Moreover, it is slightly more likely to be chosen for a junction crossing than for a mid-block crossing. Skipping one crossing opportunity was found to affect the probability of crossing at the next crossing opportunity.
- Effect of road type on mid-block crossing utility (B0_majorroad, B0_secondaryroad, B0_minorroad): secondary roads and minor roads are more likely to be chosen for mid-block crossings than major roads.
- Effect of traffic on mid-block crossing utility (B0_trafficempty, B0_trafficlow, B0_traffichighcong): pedestrians are more likely to cross at mid-block when traffic is low, and even more likely when there is no traffic, compared to when traffic is high or at congestion.
- Effect of traffic signal (B1_signal): traffic signal was found to increase the probability for junction crossing.
- Effect of barriers (B1_barriers): the presence of barriers increases the probability of crossing at junction.
- Effect of pedestrian speed (B0_speed): this effect is significant at 85%, and indicates a weak tendency of faster pedestrians to cross at mid-block.

As a next step, the components of pedestrian attitudes, perception and behavior were introduced in the model as alternative-specific explanatory variables affecting the choice of mid-block crossing. The best-fitting final model is presented in Table 3. By introducing these variables, the effect of pedestrian walking speed and the skipping of one crossing opportunity become non significant. The effects of human factors that were found significant at 90% concern component 1 and component 3:

- „Risk-taking pedestrians and optimisers" (B0_comp1) are more likely than others to cross at mid-block, as would be expected.
- „Pedestrians for pleasure‟ (B0_comp3) are (marginally) correlated with reduced probability for crossing at mid-block, in contrast to those of component 1.

As a last step of the analysis, it is examined whether there are other effects due to differences between individuals (and not due to differences between alternatives), that were not captured by the questionnaire and the resulting components identified.

These are introduced as a common random intercept of the mid-block and junction alternatives, i.e. these are allowed to randomly vary across individuals. A panel effect is therefore introduced in the model, with mean equal to zero, standard deviation equal to „sigma‟ and variance equal to „zero_sigma‟ (see Table 3). The variance of the random effect is marginally significant. Moreover, it does not appear to contribute to a remarkable improvement of model‟s fit.

It is also noted that interactions of component scores and other variables (namely road type and traffic) were tested, but were not found to be significant.

Table 3 to be inserted here

5. Discussion

The results presented above suggest that human factors are statistically singificant explanatory variables of pedestrian crossing choices. However, their contribution to the overall fit of the model is rather small. In particular, the Likelihood Ratio was slightly improved by including the human factors, but the adjusted rho-squared did not change substantially. It appears therefore that the combined analysis of pedestrians observed choices and the underlying human factors is not straightforward.

It should be kept in mind, however, that the method implemented in this research for analysing the effect of human factors on pedestrian crossing choices, is an intermediate step towards this purpose. More specifically, the successive steps for modelling pedestrian choices with human factors can be summarised as follows (Ben-Akiva et al., 1999; Ben-Akiva et al., 2002):

- Standard discrete choice models without any latent variable;
- Observed variables of human factors (i.e. questionnaire questions or indicators) may be directly inserted in the choice model; however they are highly correlated and they are not causal.
- A two-stage approach can be implemented: a principal component analysis to estimate the latent variables "components", and their scores are then introduced in the choice model. Given that only the mean component scores are introduced, however (i.e. their variance is not included), some measurement errors and inconsistent estimates may be obtained.
- Integrated choice and latent variables models can be estimated.

Obviously, in the present research the third option was explored, and it appears that the development of integrated models is an appropriate field for further research.

6. Conclusion

The results of the present analysis, based on an sample of 74 pedestrians having participated in the field survey scenarios and filled-in the questionnaire, are encouraging. The basic research hypotheses appear to be largely confirmed. More specifically, pedestrian crossing behaviour is affected by road type, traffic conditions, traffic control and pedestrian characteristics (observed or unobserved).

The implementation of Categorical Principal Component Analysis leads to the identification of 3 meaningful components of human factors of pedestrian crossing behaviour. These reveal correlations between travel motivations, risk-taking, risk perception and opinion on drivers, and some interesting patterns, such as the fact that more frequent walking is strongly correlated with less risk-taking; on the other hand, more risk-taking is correlated with low delay acceptance while walking („optimisers‟).

On the basis of the field survey data, mixed sequential logit models were developed for the probability to cross at junction, at mid-block or not cross at all on each road link of the pedestrian trip. Statistically significant parameters include road type, traffic flow and traffic control, with sign and magnitude largely in accordance with assumed in the research hypotheses.

Human factors components were introduced in the models, as fixed effects and as a random effect reflecting unobserved heterogeneity between individuals. Both fixed and random effects of human factors were found to be significant, although not strongly. The fixed effect in particular reveal a positive relationship between more self-reported risk-taking behaviour and observed mid-block crossing probability. However, the overall fit of the model was not considerably improved.

These results will be used as a basis for the development, in the next stages of the research, of „integrated choice and latent variables models‟. In these models, components of human factors (latent variables) are built and estimated within the choice model development, therefore not only taking into account the mean component scores but also the variance of the components scores.

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 address special thanks to Jean-Michel Auberlet and Marie-Axelle Granié from IFSTTAR, France, for their useful comments and suggestions in earlier stages of this research.

References

Baltes M., Chu X. Pedestrian level of service for mid-block street crossings. In the Proceedings of the 81st TRB Annual Meeting, CD-ROM, Transportation Research Board, Washington, 2002.

Ben-Akiva, M., Lerman, S.R. Discrete Choice Analysis: Theory and Applications to Travel Demand. The MIT Press, Cambridge Massachusetts, London England, 1985.

Ben-Akiva, M., Walker, J., Bernardino, A., Gopinath, D., Morikawa, T., and Polydoropoulou, A. Integration of Choice and Latent Variable Models. In: Mahmassani, H. (Ed). Perpetual Motion: Travel Behaviour Research Opportunities and Application Challenges. Elsevier Science, May 2002.

Ben-Akiva, M., McFadden, D., Gärling, T., Gopinath, D., Bolduc, D., Borsch-Supan, A., Delquié, Ph., Larichev, O., Morikawa, T., Polydoropoulou, A. and Rao, V. Extended Framework for Modeling Choice Behavior. Marketing Letters Vol. 10 No 3, 1999, pp. 187-203.

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

Bierlaire, M. (2003). BIOGEME: A free package for the estimation of discrete choice models, Proceedings of the 3rd Swiss Transportation Research Conference, Ascona, Switzerland.

Chu X., Guttenplan, M., Baltes M. Why People Cross Where They Do - The Role of the Street Environment. Transportation Research Record No 1878, 2003, pp. 3-10.

Das S., Manski C.F., Manuszak M. (2005). Walk or Wait? An Empirical Analysis of Street Crossing Decisions. Journal of Applied Econometrics 20 (4), pp. 445-577.

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

Evans D., Norman P. Understanding pedestrians‟ road crossing decisions: an application of the theory of planned behaviour. Health Education Research Vol.13, No.4, 1998, pp. 481-489.

Granié M.A., Pannetier M., Guého L. Developing a self-reporting method to measure pedestrian behaviors at all ages. Accident Analysis and Prevention Vol. 50, 2013, pp 830– 839.

Grayson G. B. Pedestrian Risk in Crossing Roads: West London Revisited. Traffic Engineering and Control Vol.28, 1987, pp. 27-30.

Hamed M.M. Analysis of pedestrians' behaviour at pedestrian crossings. Safety Science Vol.38, 2001, pp. 63-82.

Himanen, V. and Kulmala, R. An application of logit models in analysing the behaviour of pedestrians and car drivers on pedestrian crossings. Accident Analysis & Prevention Vol.20, No.3, 1988, pp. 187-197.

Honoré B.E., Kyriazidou E. Panel Data Discrete Choice Models with Lagged Dependent Variables. Econometrica Vol. 68 No 4, 2000, pp. 839-874.

Lassarre S., Papadimitriou E., Golias J., Yannis G. Measuring accident risk exposure for pedestrians in different micro-environments. Accident Analysis & Prevention Vol.39, No.6, 2007, pp. 1226-1238.

Linting, M., Meulman, J.J., Groenen, P.J.F., van der Koojj, A.J. Nonlinear principal components analysis: Introduction and application. Psychological Methods Vol. 12 No 3, 2007, pp. 336-358.

Meulman J.J, van Der Kooij A.J., HeiserKaplan W.J. 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, 2004.

Muraleetharan, T., Takeo A., Toru H., Kagaya S., Kawamura S. Method to Determine Overall Level-of-Service of Pedestrians on Sidewalks and Crosswalks based on Total Utility Value. In the Proceedings of the 83rd TRB Annual Meeting, CD-ROM, Transportation Research Board, Washington, 2004.

Oxley, J., Fildes, B., Ihsen, E., Charlton, J., and Days, R. Crossing roads safely: An experimental study of age differences in gap selection by pedestrians. Accident Analysis & Prevention Vol.37, 2005, pp. 962–971.

Papadimitriou E., Yannis G., Golias J. Analysis of pedestrian exposure to risk in relation to crossing behavior. Transportation Research Record No 2299, 2012, pp. 79-90.

Papadimitriou E.. Theory and models of pedestrian crossing behaviour along urban trips. Transportation Research Part F Vol. 15 No 1, 2012, pp. 75-94.

Sarkar, S. Evaluation of safety for pedestrians at macro- and microlevels in urban areas. Transportation Research Record No 1502, 1995.

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

Sun D., Ukkusuri S.V.S.K., Benekohal R.F., Waller S.T. Modeling of Motorist-Pedestrian Interaction at Uncontrolled Mid-block Crosswalks. In the Proceedings of the 82nd TRB Annual Meeting, CD-ROM, Transportation Research Board, Washington, 2003.

Wooldridge, J.M. Simple solutions to the initial conditions problems in dynamic, non linear panel data models with unobserved heterogeneity. Journal of Applied Econometrics Vol. 20, 2005, pp. 39-54.

Yagil, D. Beliefs, motives and situational factors related to pedestrians‟ selfreported behavior at signal-controlled crossings. Transportation Research Part F Vol. 3 No 1, 2000, pp. 1–13.

Table 1. Survey questionnaire

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)

Table 2. Component loadings for the optimally scaled questionnaire variables

Component 1: Risk taker & optimizer	
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
Component 2: Conservative & public transport user	
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
Component 3: Pedestrian for pleasure	
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

Table 3. Parameter estimates of the mixed sequential logit model

Utility functions				
0 (cross at mid-block)	=	ASC0 * one + B0_first * first + B0_majorroad * majorroad + B0_secondaryroad * secondaryroad + B0_minorroad * minorroad + B0_trafficempty * trafficempty + B0_trafficlow * trafficlow + B0_traffichighcong * traffichighcong + B0_comp1 * Comp1 + B0_comp3 * Comp3 + ZERO [SIGMA] * one		
1 (cross at junction)	=	ASC1 * one + B1_first * first + B1_signal * L_signal + B1_barriers * L_barriers + ZERO [SIGMA] * one		
2 (no crossing)	=	ASC2 * one		
Utility parameters				
Name	**Value**	**Std. error**	**t-test**	**P-value**
ASC0	-3.890	0.457	-8.510	0.000
ASC1	-2.040	0.230	-8.880	0.000
ASC2	0.000	--fixed--		
B0_comp1	0.201	0.107	1.880	0.060
B0_comp3	-0.161	0.114	-1.410	0.160
B0_first	0.893	0.252	3.550	0.000
B0_majorroad	0.000	--fixed--		
B0_minorroad	0.631	0.300	2.100	0.040
B0_secondaryroad	1.630	0.374	4.370	0.000
B0_trafficempty	1.360	0.395	3.450	0.000
B0_traffichighcong	0.000	--fixed--		
B0_trafficlow	0.664	0.317	2.100	0.040
B1_barriers	0.936	0.205	4.570	0.000
B1_first	0.978	0.206	4.750	0.000
B1_signal	0.177	0.177	1.000	0.320
SIGMA	-0.371	0.122	-3.050	0.000
ZERO		--fixed--		
Variance of normal random coefficients				
Name	**Value**	**Stderr**	**t-test**	
ZERO_SIGMA	0.138	0.104	1.32	
Model's fit				
Number of estimated parameters	13			
Number of observations	1048			
Number of individuals	74			
Number of draws	5000			
Null log-likelihood	-1043.86			
Final log-likelihood	-812.475			
Likelihood ratio test	461.223			
Adjusted rho-square	0.209			

Figure 1. Presentation of the crossing scenarios on the survey site map

Figure 2. Sequential logit model of pedestrian crossing behaviour