Journal of Environmental Psychology 82 (2022) 101807

Contents lists available at ScienceDirect

Journal of Environmental Psychology

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

Empirical characterisation of agents' spatial behaviour in pedestrian movement simulation

Gabriele Filomena [a,*], Lia Kirsch [a], Angela Schwering [a], Judith A. Verstegen [a,b]

[a] Institute for Geoinformatics, University of Münster, Germany
[b] Laboratory of Geo-Information Science and Remote Sensing, Wageningen University, Wageningen, the Netherlands

ARTICLE INFO

Handling Editor: L. McCunn

Keywords:
Empirically based agent-based modelling
Pedestrian movement
Cognitive maps
Cluster analysis
Route choice behaviour

ABSTRACT

Route choice behaviour is a key factor in determining pedestrian movement flows throughout the urban space. Agent-based modelling, a simulation paradigm that allows modelling individual behaviour mechanisms to observe the emergence of macro-level patterns, has not employed empirical data regarding route choice behaviour in cities or accommodated heterogeneity. The aim of this paper is to present an empirically based Agent-Based Model (ABM) that accounts for behavioural heterogeneity in pedestrian route choice strategies, to simulate the movement of pedestrians in cities. We designed a questionnaire to observe to what degree people employ salient urban elements (local and global landmarks, regions, and barriers) and road costs (road distance, cumulative angular change) and to empirically characterise the agent behaviour in our ABM. We hypothesised that a heterogeneous ABM configuration based on the construction of agent typologies from empirical data would portray a more plausible picture of pedestrian movement flows than a homogeneous configuration, based on the same data, or a random configuration. The city of Münster (DE) was used as a case study. From a sample of 301 subjects, we obtained six clusters that differed in relation to the role of global elements (distant landmarks, barriers, and regions) and meaningful local elements along the route. The random configuration directed the agents towards natural elements and the streets of the historical centre. The empirically based configurations resulted in lower pedestrian volumes along roads designed for cars (25% decrease) but higher concentrations along the city Promenade and the lake (40% increase); based on our knowledge, we deem these results more plausible. Minor differences were identified between the heterogeneous and homogeneous configurations. These findings indicate that the inclusion of heterogeneity does not make a difference in terms of global patterns. Yet, we demonstrated that simulation models of pedestrian movement in cities should be at least based on empirical data at the average sample-level to inform urban planners about areas prone to high volumes of pedestrians.

1. Introduction

Walking in urban spaces is a multifaceted act long discussed in transport geography (Hill, 1984; Papadimitriou et al., 2009), urban geography (Evans & Jones, 2011; Middleton, 2011), and psychological research (Alfonzo, 2005; Darker et al., 2007; Gidlow et al., 2016). Rather than a simple way to get from A to B, walking has been seen as a political act, 'a form of urban emancipation' (De Certeau, 1984), and a social practice (Middleton, 2018); it is considered a predictor of physical and mental well-being (Ferdman, 2019; Pucher & Buehler, 2010; Roe & Aspinall, 2011). Moreover, walking has become a major topic in the discourse on low-carbon emitting mobility, as a form of sustainable travel (Tight et al., 2011) that fosters liveable cities (Stratford et al., 2020). Inevitably, acquiring knowledge on how pedestrians distribute in urban spaces constitutes a fundamental step for observing pedestrian negotiation processes with the city (Willis et al., 2004), advancing theories of pedestrian behaviour (Zacharias, 2001), and thinking and designing urban spaces (Torrens, 2016).

Agent-based modelling has recently been adopted to study and project the movement of pedestrians in urban spaces at different scales. An Agent-Based Model (ABM) consists of a system of individual entities, *agents*, equipped with a set of behavioural rules that shape their decisions and interactions in and with the *environment* (Bonabeau, 2002). An 'agent senses that environment and acts on it, over time' (Franklin & Graesser, 1996, p. 25). As such, agent-based modelling 'permits one to study how rules of individual behavior give rise - or "map up" - to

* Corresponding author.
E-mail address: gabriele.filomena@uni-muenster.de (G. Filomena).

https://doi.org/10.1016/j.jenvp.2022.101807
Received 16 July 2021; Received in revised form 17 February 2022; Accepted 7 April 2022
Available online 20 April 2022

macroscopic regularities and organizations' (Epstein, 1999, p. 41). Not only has this paradigm established itself for understanding geographical and social phenomena (Crooks et al., 2008; Heppenstall et al., 2016), but it has also been embraced in environmental and social psychology to better situate human interactions in the context of social and spatial environments (Jager, 2017; Jager & Ernst, 2017).

The 'complexity and dynamic nature of the cognition behind wayfinding' processes in cities (Spiers & Maguire, 2008, p. 246) makes agent-based modelling a natural choice for analysing pedestrian behaviour in urban spaces (Kerridge et al., 2001). Simulation models of pedestrian movement in cities have focused on route choice behaviour across street networks (for a review see Papadimitriou et al., 2009; Filomena et al., 2020), on pedestrian crossing behaviour (see Papadimitriou et al., 2009), or locomotion and gait realism (e.g. Torrens, 2012). However, in transport geography and GIScience, walking behaviour has been treated as a 'homogeneous and largely self-evident means of getting from one place to another' (Middleton, 2011, p. 92). Assumptions on route choice behaviour - that is the combination of processes that allow one to find a viable route between two locations - strongly grounded on a rational and utilitarian perspective to human behaviour (Lorimer, 2010) have been hindering the conception of genuine simulation models of pedestrian movement in cities. Most of the existing ABMs (e.g. Jiang & Jia, 2011; Omer & Kaplan, 2017) contemplate the existence of a homogeneous set of agents who minimise certain road attributes, such as road distance or cumulative angular change. These approaches reduce the concept of urban form and its role in pedestrian behaviour (Handy, 1996) to the sole perception of the attributes of the street segments.

In contrast, in previous work, in light of research on the relationship between urban form and route choice behaviour (e.g. Epstein & Vass, 2014; Lynch, 1960; Mallot & Basten, 2009; Siegel & White, 1975), we have modelled the functions of meaningful urban elements in pedestrian movement simulation (landmarks: Filomena & Verstegen, 2021; regions and barriers: Filomena et al., 2020) to enrich the agent's cognitive representations of the environment and account for complex route choice processes. Nevertheless, such an enhanced authenticity was introduced by defining homogeneous groups of agents. This is not in line with a wealth of studies that have described differences between people's route choice strategies (e.g. Golledge, 1995; Guo & Loo, 2013; Kato & Takeuchi, 2003; Shatu et al., 2019).

Fifteen years ago, when agent-based modelling was surfacing as a method to advance the study of complex socio-environmental systems, Janssen and Ostrom (2006) introduced the term *empirically based agent-based model*. Thereby, the authors expressed the need for new rigorous models that could be applied at different scales of analysis and whose results could be generalised to different contexts. The authors were mainly calling for methods that would 'help confirm patterns observed in agent-based modelling' (Janssen & Ostrom, 2006), namely to validate models. Rounsevell et al. (2012) reformulated the concept of empirically based models as a framework to 'empirically ground the representation of human behavioural processes' (Rounsevell et al., 2012) in the context of socio-ecological systems. In their view, such models should account for the behavioural heterogeneity existing in the modelled population to make more genuine interactions and patterns emerge: 'the individual outcomes and interactions that result from that heterogeneity is a key quality that sets the ABM approach apart from equation-based models' (Rounsevell et al., 2012, p. 261). The authors suggest using the results of questionnaires, participant observations, and other empirical approaches in the social sciences to design or parametrise behavioural rules in the ABM.

This framework has been widely adopted in research on residential choice. For instance, Haacke et al. (2022) combined hierarchical clustering techniques and qualitative data collection to derive groups of individuals characterised by alike residential choice behaviour for the initialisation of an ABM for residential mobility. Similarly, Brown and Robinson (2006) and Fernandez et al. (2005) derived agent typologies from the responses to a questionnaire on preferences for environmental characteristics and residential locations. When validating an empirically based ABM for residential mobility with aggregated data, Buchmann et al. (2016) found that behavioural heterogeneity in the ABM population would bring about different and more realistic patterns in contrast to a homogeneous specification of the model. Crowd motion modellers, interested in mimicking and understanding phenomena such as crowd egress, congestion, and wayfinding dynamics in hospitals, shopping centres, or stations (for a review see Duives et al., 2013; Yang et al., 2020), have also moved towards data-driven approaches. In this context, Haghani (2020a,b) reviewed and discussed how laboratory research, drills, virtual reality experiments, and field observations have been deployed to shed light on crucial aspects for the design of crowd simulation models (e.g. decision-making processes, locomotion, the role of contextual factors).

Nevertheless, heterogeneity - i.e. variation in route choice strategies - is not included in existing simulation models of pedestrian movement in cities. On the one hand, as for the modelling of other phenomena, this might derive from the not easy formalisation of a clear set of behaviours into algorithms, due to different pieces of empirical evidence and data (Jager & Ernst, 2017; Janssen & Ostrom, 2006). On the other hand, research on route choice strategies, although not scarce, usually refers to vehicular traffic and general wayfinding behaviour (e.g. Golledge, 1995; Jan et al., 2000; Zhu & Levinson, 2015), or aims at identifying to what extent pedestrian routes diverge from the shortest path (e.g. Foltête & Piombini, 2010; Guo & Loo, 2013; Koh & Wong, 2013; Yang & Diez-Roux, 2012). The latter results are not straightforward to generalise into parameters in an ABM, as they often refer to specific environmental factors (e.g. presence of green areas, pedestrian facilities, etc.).

The aim of this paper is to present an empirically based ABM (Rounsevell et al., 2012) for the simulation of pedestrian movement in urban spaces that accounts for behavioural heterogeneity in pedestrian route choice strategies. Walking is modelled as a situated behaviour resulting from continuous and mutual interactions between an individual's cognitive system and the perceived environment. We describe an ABM in which the agent's behavioural parameterisation, which regulates the effect of certain urban elements on the agent's route choice strategies, is empirically characterised. A questionnaire investigating people's route choice strategies in pedestrian movement is used to: a) empirically support the theoretical assumptions underlying the structure of the agents' cognitive maps and the ensuing behaviours; b) derive groups (clusters) of subjects featured by similar behavioural properties and incorporate them into the ABM by defining agent typologies; c) set ranges of values for global and typology-based model parameters, i.e. to calibrate the ABM. Hence, in this work, we address the following research questions: a) What is the diversity in route choice strategies as concerns the usage of minimisation heuristics and information about meaningful urban elements? b) To what extent does the variation in the agents' heterogeneity - i.e. a model including agent typologies *vs* a model with a homogeneous set of agents - generate different movement patterns across the street network? The city of Münster (DE) is used as a case study.

The paper is organised as follows: in the methodology section, we present the framework that we embraced to empirically characterise the ABM through a questionnaire. Firstly, the general functioning of the model and its parameters are introduced. Secondly, the study's participants and the structure of the questionnaire are described. Finally, the cluster analysis approach employed to derive groups of individuals and the procedures to define the agent typologies are outlined. In the results section, we present the findings of the study and the clustering; furthermore, we contrast and discuss the movement patterns generated from the different ABM configurations.

2. Materials and methods

2.1. Overview

In previous work, we have stressed the importance of including cognitive representations of space - *cognitive maps* - in the architecture of pedestrian agents. This was achieved by enriching the geographic environment with meaningful urban elements (UE) - landmarks, nodes, regions, and barriers, following Kevin Lynch's metaphor of the Image of the City (Lynch, 1960) - that shape the development of cognitive maps. In earlier versions of our simulation model, agents were equipped with route choice approaches that reshaped spatial decisions in light of: the identification of on-route marks (local landmarks) and distant orienting landmarks (global landmarks) (see Filomena & Verstegen, 2021); the perception of urban subdivisions as a consequence of regionalisation processes and the presence of barriers (see Filomena et al., 2020). In contrast to our conceptualisation, existing ABMs have mainly depicted agents who exclusively make use of road costs for formulating routes, i.e. by minimising a certain cost at the global city level (GMH).

Therefore, in our ABM, agents could use different route choice models based on meaningful urban elements and road attributes employed during *prospective planning* (or coarse plan; see Wiener & Mallot, 2003):

- (GMH) Global road cost minimisation heuristics, i.e. road distance, cumulative angular change.
- (UE) Regions (identification of gateways).

and during *situated planning* (or fine planning), a readjustment and refinement navigation phase:

- (LMH) Local road cost minimisation heuristics, i.e. road distance, cumulative angular change.
- (UE) On-route marks.
- (UE) Orienting distant landmarks.
- (UE) Natural and severing barriers.

In this study, we present an enhancement of the model that contemplates the concurrent usage of multiple urban elements and allows for different route choice strategies within the pedestrian agent population (see below, Section 2.2). We developed a questionnaire to empirically ground and inform the mechanisms of our ABM, as well as to calibrate the parameters that regulate the interplay of such urban elements throughout the agent's route formulation process (Section 2.3). The results of the questionnaire were used to derive groups (clusters) of individuals who shared similar route choice strategies and consequently incorporated into the ABM to build agent typologies (Section 2.4). The model was evaluated by comparing the distribution of pedestrian agents across the street network resulting from three different ABM configurations (Section 2.6):

1. The *null configuration*; herein, the agent population is homogeneous and the behaviour is not informed by empirical data. Rather, the agent behaviour is regulated by parameters whose values are drawn from uniform distributions between the minimum and maximum values of the parameters.
2. The *homogeneous configuration*; it is characterised by a homogeneous set of agents whose parameter values are, however, regulated by the mean and standard deviation values of different variables obtained from the responses to the questionnaire, over the whole study sample.
3. The *heterogeneous configuration*; in this configuration, the results of the cluster analysis - the number of clusters, their portion of subjects over the entire study sample, and attribute values (the mean and standard deviation of the variables) - are incorporated into the ABM to regulate the agent behaviour; the agent typologies are designed so that, when an agent is generated, it inherits its behavioural parameter values from the corresponding cluster attribute values.

2.2. The ABM and the behavioural mechanisms

The core functioning of the ABM consists of a number (A) of pedestrian agents that generate trips between different pairs of locations - OD pairs - across an urban environment, on the basis of their route choice strategies (Fig. 1). The urban environment includes the street network, the functional and physical attributes of the buildings, natural and artificial elements (natural and severing barriers). Such an environment informs the agent's cognitive map of the city through the incorporation into the street network of landmark scores, the membership of the street segments to regions, their proximity to barriers, and road cost attributes (Filomena et al., 2019). All agents are endowed with a cognitive map of the environment, but may rely on different pieces of information for formulating their routes. At the end of a single model execution, the ABM stores the number of times a street segment was crossed by the agents, thus computing *pedestrian volumes* per segment. The model was implemented in GeoMASON, a multi-agent simulation environment written in the Java programming language (Sullivan et al., 2010).

Each agent is characterised by different *ABM parameter* values, expressed as probabilities, that regulate its reliance on the urban elements and road attributes represented in the cognitive map (Table 1). As compared to existing models, the behavioural mechanisms of the agent - i.e. the way it uses urban information to complete a path - are not static: every time the agent formulates a path, it may be using a different set of elements and road attributes, depending on the agent's characteristics and the type of route (e.g. length, complexity, etc.). The definition of the route choice behaviour of the agent is structured hierarchically through the prospective and situated planning phases. It entails different steps, regulated by *stochastic discrete parameters* (orange boxes in Fig. 1, see also Table 1):

1. Cost minimisation only (p_c) *vs* usage of urban elements (p_e). If an agent chooses to only minimise costs the other steps are not considered; a GMH is instead picked (p_d *vs* p_a). If the agent decides to make use of the UE, the next steps are considered.
2. Usage of regions (p_r), or not; the sequence of regions is however generated when the origin and the destination are at least x meters away from each other.
3. Choice of the local minimisation heuristic to adopt: road distance or cumulative angular change (p_{ld} *vs* p_{la}).
4. Usage of sub-goals: on-route marks (p_o), barrier sub-goals (p_b), or none.
5. Usage of distant landmarks (p_g), or not.

Where p_e and p_c are complementary and sum to 1.0, as well as p_d and p_a, and p_{ld} and p_{la}. The sum of all values of the UE parameters (p_r, p_o, p_b, p_g) may be higher than 1.0 to account for a possible simultaneous usage of the urban elements, but each of them sums to 1.0 with its corresponding complementary parameter value (e.g. using regions *vs* not using regions).

Furthermore, *stochastic preference parameters* (green boxes in Fig. 1) direct the perception of costs of street segments located near barriers:

- Preference for segments along or within natural barriers, μ_n and σ_n.
- Aversion to road segments that constitute, extend, or cross severing barriers, μ_s and σ_s.

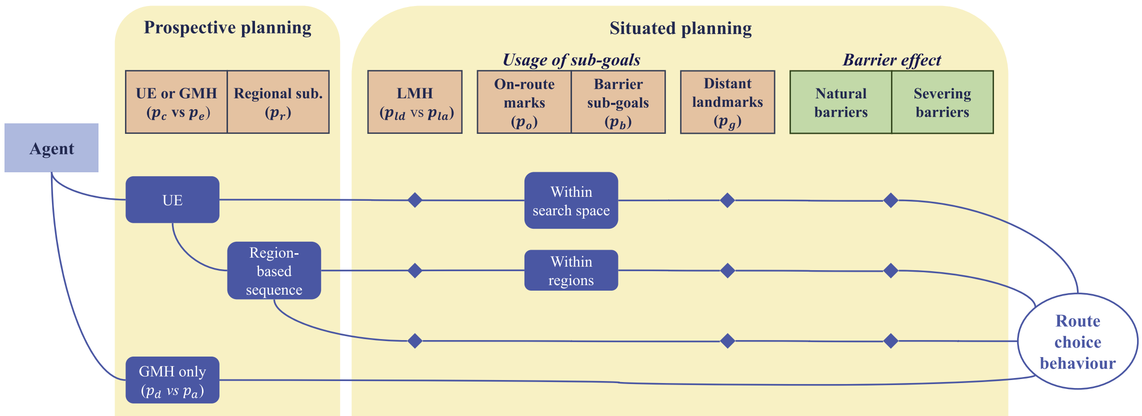

Fig. 1. The definition process of the agent's route behaviour and the parameters involved, at different stages. Diamonds indicate the possible activation of a preference or reliance on an urban element.

Table 1
The behavioural components of the ABM and the corresponding parameters that regulate the definition of the agent's route choice behaviour. They fall within *PP*: prospective planning and *SP*: situated planning. The ABM parameters represent probabilities and they can assume values between 0.0 and 1.0. In the empirically based configurations of the ABM, the sample attribute values are used to obtain the ABM parameter values (see Eqs. (5) and (1)). *Sample attributes* may be *cluster attributes* when agent typologies are built.

Phase	Behavioural component	ABM parameter	Sample attributes
PP	Completing a route exclusively minimising road costs	p_c	μ_c, σ_c
	Completing a route making use of urban elements	p_e	μ_e, σ_e
	Completing a route exclusively minimising road distance	p_d	μ_d, σ_d
	Completing a route exclusively minimising cumulative angular change	p_a	μ_a, σ_a
	Formulating a region-based sequence	p_r	μ_r, σ_r
SP	Locally minimising road distance	p_{ld}	μ_{ld}, σ_{ld}
	Locally minimising cumulative angular change	p_{la}	μ_{la}, σ_{la}
	Making use of on-route marks	p_o	μ_o, σ_o
	Making use of barrier sub-goals	p_b	μ_b, σ_b
	Relying on distant landmarks	p_g	μ_g, σ_g
	Preference for segments along or within natural barriers	μ_n, σ_n	μ_n, σ_n
	Aversion to road segments that constitute, extend or cross severing barriers	μ_s, σ_s	μ_s, σ_s

As they represent natural and restorative elements of the urban environment, the first type of barriers attracts pedestrians and may induce them to take detours. Severing barriers, on the contrary, discourage pedestrians from approaching them, as they are difficult to cross, unsafe, or unpleasant. The perceived cost i_e of a street segment e is modelled in the ABM as:

$$i_e = cost_e * Z_d$$
$$with\, Z_d \sim \begin{cases} min(N(\mu_n, \sigma_n), 1.0), & \text{if } e \text{ lies along or within a natural barrier} \\ max(N(\mu_s, \sigma_s), 1.0), & \text{if } e \text{ is, crosses, or lies along a sev. barrier} \\ N(1.0, 0.10), & \text{in any other case} \end{cases} \tag{1}$$

where $cost_e$ is the actual cost (e.g. road distance) of the segment; Z_d is a normal distribution with a mean μ_n and a standard deviation σ_n, when the segment lies along natural barriers, a mean μ_s and a standard deviation σ_s when it is placed in the proximity of severing barriers, or a mean 1.0 and a standard deviation 0.10 in all other cases.

Table 2
The socio-demographic profile of the participants (N = 301). The variables describing the relationship with the city should not be considered mutually exclusive.

	Demographics	Frequency	%
Gender	Female	188	62.5%
	Male	110	36.5%
	Non-binary	2	0.6%
	Prefer not to say	1	0.4%
Age group	18–25	147	48.8%
	26–33	114	37.9%
	34–41	15	5.0%
	42–49	7	2.3%
	50–57	8	2.7%
	58–65	8	2.7%
	66–73	2	0.6%
Relationship with the city	(used to) live	200	66.4%
	(used to) work	93	30.9%
	(used to) study	255	84.7%
	Occasional visitor	16	5.3%
	Tourist	1	0.3%

In summary, an agent firstly "decides" whether to use only minimisation approaches based on road costs or resort to urban elements. In the latter case, a further decision is taken as regards the elaboration of a sequence of regions. Afterwards, a local minimisation heuristic is chosen. In the situated planning phase, the complexity of the environment comes into play: the agent may segment the route in further chunks by identifying sub-goals such as on-route marks (a complex environment may require more on-route marks than a more legible space) or barrier sub-goals. Finally, the fine choice of certain roads, while influenced by the LMH adopted by the agent, can be shaped by the need to keep in sight distant landmarks towards the destination, a preference for roads that extend along natural barriers, or an aversion to roads adjacent to severing barriers. As a result, the final route could range from a purely cost-based route to a path emerging from the perception and usage of several urban elements.

2.3. The study

2.3.1. Participants

The participants were recruited from across the student population and members of local associations interested in urban mobility or citizen participation. The completion of the questionnaire was rewarded with a 10 euro bank transfer. From an original sample of 418 subjects, we disregarded the records of the participants who took less than 20 min to complete the questionnaire. Thus, the final sample consisted of 301 participants, 188 women and 110 men (Table 2). Age ranged from 18 to 70 years, with a mean of 27.8 and a standard deviation of 9.3 years. On average, participants had lived in the case study area for 8.1 years (standard deviation: 10.8 years); this variable was used as a proxy for the degree of *familiarity* with the case study area.

2.3.2. Materials

An online questionnaire was designed to investigate pedestrian route choice strategies and preferences for certain properties of the environment. The Limesurvey[1] online tool was used to build a questionnaire composed of six sections. The use of a mouse device, a touch pad, or a touchscreen display, in combination with a keyboard, was necessary to complete the questionnaire.

In Section I, a set of questions was devised to collect information about the walking habits of the participants (e.g. the number of days per week they would engage in walking trips in the city, type of destinations, etc.). Sections II, III, and IV, the core of the questionnaire, consisted of three navigational video tasks. Videos that simulate a walking experience have been used in other studies for different purposes (e.g. Bornioli et al., 2018; Gatersleben & Andrews, 2013), as an alternative to the actual act of walking in the urban environment. For these three tasks, we generated a set of routes between three pairs of locations by means of our ABM - *model-generated routes* (see for example Fig. 2, left panel) - to maintain a correspondence between the questionnaire and the ABM. The Euclidean distance between the three OD pairs was 1234, 1698, and 1792 m, respectively. The model-generated routes were obtained through route choice models that combined the usage of one urban element (UE) with a local minimisation heuristic (LMH) or exclusively minimised road costs (GMH) (see above). Thus, we identified a set of decision points, namely street junctions where the model-generated routes would intersect. After having derived the routes from the model on a graph representation of the street network, we traversed them with a camera and recorded the corresponding urban scenes, between each decision point. The resulting videos were embedded in the questionnaire.

Section V, which aimed at collecting preferences for certain route characteristic, included three digital maps of sub-areas of the case study area; each map displayed three different routes between an origin and a destination, computed on the basis of road distance, least cumulative angular change, and fewest intersections shortest paths, respectively. Finally, in Section VI, demographic information - i.e. age, gender, belonging to a certain category (student, professional, tourist, or occasional visitor), time lived in the case study area, and district of residence (if applicable) - was collected.

The responses of the subjects were automatically recorded in a table generated by the online tool.

2.3.3. Procedure

The study was approved by the Institutional Review Board of the University of Münster. The questionnaire could be accessed via a Web link. The questionnaire was estimated to take 30–60 min. The participants were told that the study investigated preferences regarding their walking behaviour in the case study area. After agreeing to participate, the subject would go through six different sections. In Section I, the participants indicated how many days per week they would walk, how often they would walk for a certain reason (on a scale from 'never' to 'always'), and to what extent they resorted to navigation devices when walking. They were also asked to express their degree of agreement with four statements extracted from a validated questionnaire on sense of orientation and spatial skills (see Münzer et al., 2016).

At the beginning of Sections II, III, and IV, each asking the subjects to virtually complete a route between an origin and a destination, an initial 180° scene from the origin of the route was shown. The participants were asked to choose in which direction they would like to head; following their choice, the corresponding video was reproduced to mimic the walking experience along that section of the route till a decision point. From there, subjects were again asked to choose in which direction they would continue their route (e.g. go straight, turn left, etc.; see Fig. 2, right panel); then, they were shown the following videos, through the next decision points, up to the destination. At the end of each section, the subjects were told to report their cognitive load to verify whether the tasks caused fatigue and mental effort. The Paas subjective cognitive load scale (Paas, 1992) was used for this purpose.

In Section V, the participants were shown three different maps. For each map, the subjects were asked to pick the route that they would be more likely to walk amongst the ones displayed. Furthermore, in Question 4, subjects were asked to express 'How important are the following properties of a route if you have to reach a certain place by foot?':

1. The route traverses or extends along green areas (e.g. parks, gardens, forests).
2. The route extends along water bodies (e.g. rivers, lakes).
3. The route does not cross or go along major roads, such as motorways, national roads, or multilane roads.
4. The route features streets with wide sidewalks or for exclusive pedestrian use.
5. The route crosses safe areas (e.g. good lighting, presence of other people, the area is known).
6. The route allows experiencing nice scenic views (e.g. natural landscapes, buildings of historical and architectural interest).
7. The route crosses interesting or lively districts.

The responses to Question 4 were to be given on the scale 'not important', 'slightly important', 'important', 'fairly important', 'very important'. Finally, the participants would fill in Section VI providing demographic information.

Two pretest phases were conducted. Initially, a partly different questionnaire had been designed; this would mainly include route description tasks and route choices on maps. However, the first type of task appeared to require a high cognitive load on the Paas scale (Paas, 1992) for the tested participants (N = 10), and was not consistent with the aim of the study (see Hölscher et al., 2011, for the divergence between the route that people would describe and the route they would walk between two locations). Therefore, we introduced the video tasks to induce a more situated experience for the participant. A second test phase was carried out for appraising the reviewed version of the questionnaire; no relevant remarks were made and the participants (N = 5) judged the questionnaire enjoyable to complete.

2.4. Identifying groups of similarly behaving individuals

Cluster analysis identifies groups of objects, individuals, or entities that exhibit comparable properties. For several disciplines, describing certain entities or subjects by means of labels allows summarising characteristics across a set of observations (Everitt et al., 2011). In this work, as the input of the cluster analysis, we used variables, expressed in

[1] https://www.limesurvey.org.

Fig. 2. *Left*: Sections of the model-generated routes (some may be part of more than one route choice model) for one of the three OD pairs and the derived decision points - i.e. junctions where the participants were asked to choose amongst two or more alternatives towards the destination. The map is oriented north. *Right*: An example of a decision point during the virtual navigation. In this case, the participants were asked whether they would like to proceed by taking the path on the right or going left, down the underpass. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

probabilities, obtained from the portions of each participant's routes (Sections II, III, and IV of the questionnaire) that overlap with the model-generated routes. The variables are listed in Table 1 and represent the behavioural components modelled in the ABM by means of the stochastic discrete parameters. A logarithmic transformation of these variables was performed to reduce the skewness exhibited by the distribution of the raw variables. Additionally, a *z-scores* standardisation was performed on the transformed values (Steinley, 2004).

The features 'completing a route exclusively minimising road costs' and 'completing a route making use of urban elements' were derived from the intersection of the GMH and UE variables, respectively, and therefore not employed for identifying the clusters. The responses to Question 4 in Section V were converted into a scale from 0.0 to 1.0. Afterwards, the variable 'preference for segments along or within natural barriers' was obtained as the mean of the responses to sub-questions 1 and 2, Question 4; the variable 'aversion to road segments that constitute, extend, or cross severing barriers', on the grounds of the responses to sub-question 3, Question 4. They were also not included in the cluster analysis.

We determined the sample size by adopting the approach advanced by Formann (1984), according to which, for cluster analysis, the sample should include at least 2^u subjects or observations, where u represents the number of variables. Given that u is equal to 8 variables in this work, the minimum sample size was 256.

We carried out a cluster analysis to identify similar groups of pedestrians in the study sample and build the corresponding agent typologies in the ABM. We employed the *k-means* clustering method. K-means (Lloyd, 1982) is based on the concept of cluster centroids and seeks to identify centroids that optimise intra-cluster homogeneity. The algorithm starts by identifying representative points as initial centroids. Each other object is then assigned to the closest centroid, depending on a measure of proximity, and the centroids are updated. These steps are executed until the centroids stop changing significantly.

As the k-means algorithm requires the researcher to specify the desired number of clusters, we recursively ran k-means with different desired clusters, from 3 to 9. To identify the best partition emerging from these iterations and a suitable number of clusters, a subjective examination of the clustered structures and their attributes was accompanied by the computation of the silhouette coefficient (Rousseeuw, 1987), an intrinsic measure of clustering goodness, and the Variance Ratio Criterion (VRC) (Caliński & Harabasz, 1974) coefficient ω, an index designed to determine the optimal number of clusters (Milligan & Cooper, 1985). The silhouette score considers both intra- and inter-cluster distances and it is computed for a certain partition as the average silhouette coefficient of all observations:

$$S = \frac{\sum_{x=1}^{N} \frac{b_x - a_x}{max(a_x, b_x)}}{N} \tag{2}$$

where N is the number of observations, a_x indicates the average distance from x to all other entities in x's cluster, and b_x the average distance from x to all the other points not belonging to x's cluster. The coefficient can range between -1 (objects are closer to other clusters' objects than to their own cluster's members) and 1 (compact clusters, well separated from others) (Han et al., 2012).

The VRC score of a partition is computed as:

$$VRC = \frac{SS_B}{K-1} \bigg/ \frac{SS_W}{N-K} \tag{3}$$

where SS_B is the between-cluster sum-of-squares, SS_W the within-cluster sum-of-squares, K the number of clusters, and N the number of observations. The VRC is also called *pseudo F* as it corresponds to the *F value* resulting from a one-way ANOVA, with K number of factors. Small values of SS_W and large values of SS_B, and therefore high VRC score values, indicate a partition that features well-separated clusters. Since a large number of clusters may result in lower VRC score values, it is suggested to compute the differences in the VRC scores obtained by consecutive executions of the k-means algorithm (coefficient ω) and select the lowest value (indicating a relative increase in VRC scores).

Hence, for an iteration with K desired clusters, we calculated the coefficient ω_K as:

$$\omega_K = (VRC_{K+1} - VRC_K) - (VRC_K - VRC_{K-1}) \tag{4}$$

where VRC_K is the VRC score of the partition resulting from the k-means with K desired clusters; VRC_{K+1} and VRC_{K-1} represent the scores of the previous and successive iterations.

2.5. ABM calibration

In the null configuration, the values of the stochastic discrete parameters listed above (see Section 2.1, Table 1) are considered to be uninformed and thus vary over the agent population based on a standard uniform distribution. On the contrary, while in the homogeneous configuration the behavioural parameter values are derived from the corresponding *sample attributes*, in the heterogeneous configuration an agent inherits such parameter values from the attributes of its cluster (*cluster attributes*); a sample attribute indicates the mean and standard deviation values of a certain variable over the entire study sample, a cluster attribute refers to the mean and standard deviation values of a variable within a cluster. In detail, in the empirically based configurations, the value of an agent's stochastic discrete parameter p_j (e.g. the likelihood of formulating a region-based sequence) is computed before a trip as:

$$\begin{aligned} p_j &= Z_j \\ \text{with } Z_j &\sim N(\mu_j, \sigma_j) \end{aligned} \tag{5}$$

where Z_j is a normal distribution with a mean of μ_j and a standard deviation σ_j. Hence, μ_j, and σ_j represent the mean and standard deviation values of a sample attribute in the homogeneous configuration or of a cluster attribute in the heterogeneous configuration.

In the null configuration, the stochastic preference parameters (see Section 2.1, Table 1) are also considered to be uninformed: μ_n and μ_s are drawn from a standard uniform distribution and input into Eq. (1) when running the ABM; σ_n and σ_s are set to 0.10. In the homogeneous and heterogeneous configurations, the mean and standard deviation values of these two sample or cluster attributes are directly input into Eq. (1) to model the perception of road costs during the agent navigation. In all configurations, the input value of the parameter 'aversion to road segments that constitute, extend, or cross severing barriers' is shifted by one unity, thus rescaled into the range 1.00–2.00.

2.6. ABM evaluation

The three configurations (see Section 2.1) were executed T times each as Monte Carlo simulations to balance the randomness entailed by the selection of the OD pairs and the stochastic functions of the ABM. For each configuration, the *pedestrian volumes* were determined from the pedestrian counts of the street segments (median value over the T model executions). To visualise the distribution of the pedestrian agents, we generated a figure representing the pedestrian volumes for the null configuration across the entire street network. Furthermore, for each segment, we verified whether the frequency distribution of the pedestrian volumes over the T executions in the homogeneous and heterogeneous configurations differed significantly from the null configuration. The Wilcoxon test (Wilcoxon, 1945), a non-parametric version of the *t-test*, was used for this purpose with a 0.05 α *value*. Thereby, we obtained a second figure indicating for which segments the empirically based configurations generated statistically significant different pedestrian volumes from the null configuration.

2.7. The case study area and the ABM general parameters

The study was conducted in Münster, a city situated in the north-west of Germany (North-Rhine Westphalia) with a population of around 300,000 citizens (approximately 123,000 in the central urban districts) (Münster, 2021b). The city has been named 'Germany's cycling capital' for the extent of its bike lane network, roads and spaces designed for bikes, and for the bike owning rate, equal to almost two bikes per citizen (Münster, 2021a). Whereas the city even claims to be one of the most liveable cities in the world (Münster, 2021c), the life of pedestrians is probably not as easy as the life of cyclists. Pedestrian trips had declined drastically 20 years ago, likely when the municipality encouraged a modal shift towards bicycles. Although some roads are entirely pedestrianised, pedestrians reclaim more street space, as sidewalks are often occupied by parked cars, sacrificed to bike lanes, or not large enough for people with reduced mobility (see a recent initiative by the association Münster zu Fuß, 2021). The area within a bounding box of 2500 m from the centre of Münster was used in the ABM (Fig. 3).

In the ABM, the total number of agents A was set to 301, each completing 3 trips. The number of ABM executions T was set to 20. To generate the set of OD pairs for the ABM, we first categorised each building into 'residential', 'work', or 'visit', following the categorisation introduced by Dovey and Pafka (2014). Secondly, we used the responses to Question 3 in Section I of the questionnaire to identify the likelihood of engaging in a walking trip for different purposes. Hence, the origin nodes were randomly drawn from a set of nodes whose closest building was categorised as 'residential'. The destinations were extracted amongst nodes whose closest building was categorised as:

- 'work', 30% of the time, to represent the likelihood of walking for commuting to a work or university place, 13% and 17% respectively.
- 'visit', 46% of the time, to represent the likelihood of walking for engaging in social or spare time activities, 22% and 24% respectively.

For the remaining 24% of the time, a random destination was chosen to include pedestrian trips motivated by 'other daily errands and commitments'.

3. Results

3.1. The questionnaire

On average, the participants took 53 min to complete the questionnaire. In Section I, the subjects indicated that on average they would walk 'sometimes' or 'often' for daily errands or leisure activities; less for social activities ('sometimes') or for commuting to school/university ('seldom' or 'sometimes') and work ('seldom'). Furthermore, a large number of participants stated to 'agree' or 'completely agree' with the four statements measuring the self-reported spatial skills. None of such statements taken singularly positively correlates with the subjects' familiarity with the case study area. Yet, the combination of the responses in a unique measure of spatial ability moderately correlates with familiarity (0.21[2]).

The distribution of the variables obtained from the subjects' choices in Sections II, III, and IV of the questionnaire, the video tasks, are shown in Fig. 4. Within our study sample, people's routes overlapped with the model-generated routes shaped by the urban elements in more than 60% of the cases; conversely, they followed routes based on the usage of minimisation heuristics at the global level in less than 40% of the cases. Considering that these figures take into account possible overlaps between the different routes, we can argue, in line with previous empirical evidence (e.g. Foltête & Piombini, 2010; Kim, 2015; Muraleetharan & Hagiwara, 2007), that pedestrians do not only minimise road costs when formulating a route in the urban environment. Rather, they can be

[2] Pearson product-moment correlation coefficient.

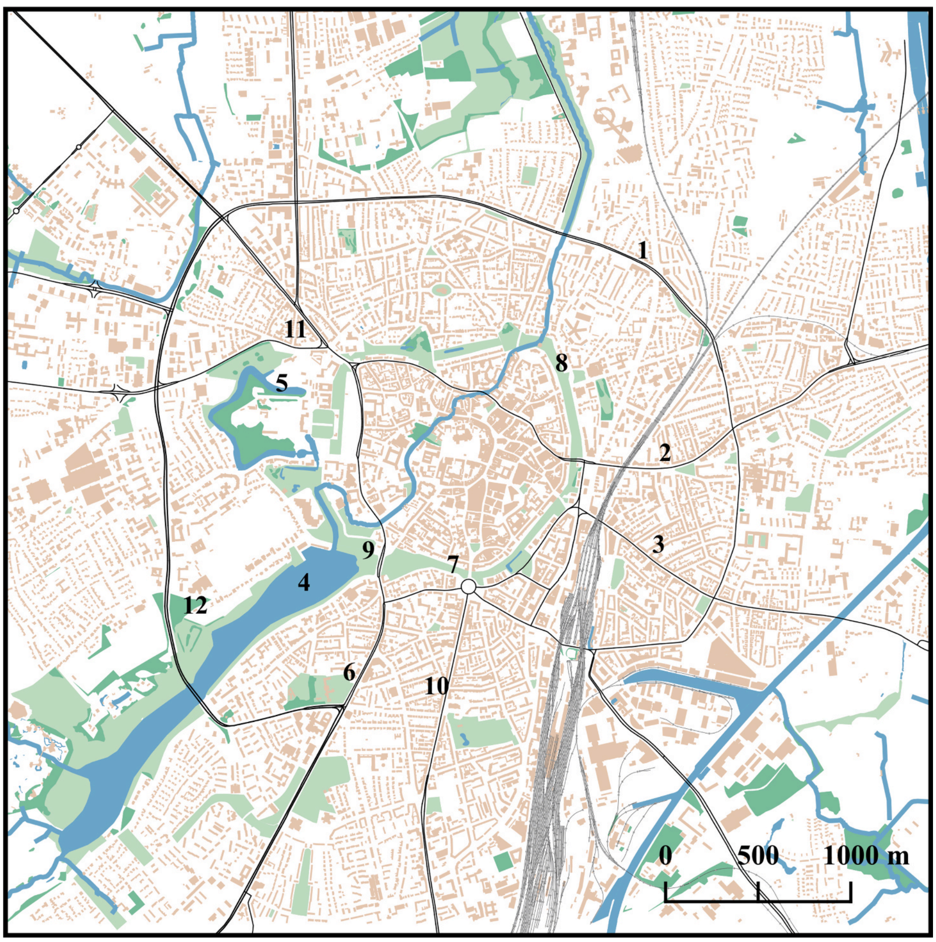

Fig. 3. The case study area: Münster, Germany (bounding box of 2500 m from the city centre). Natural elements, building footprints, main roads, and railway links within the case study area. Numbered locations are referred to in the results and discussion section. The map is oriented north. Data source: OpenStreetMap data (OpenStreetMap contributors, 2021).

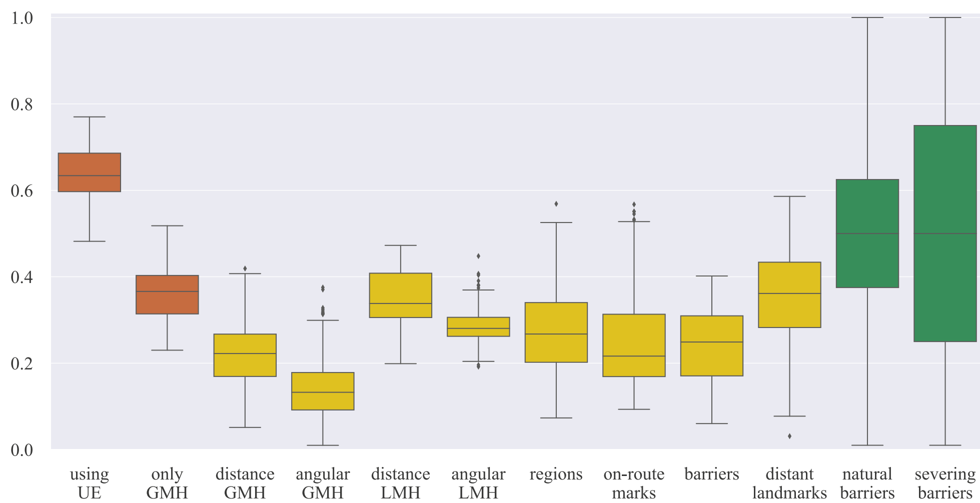

Fig. 4. Boxplot of the variables extracted from the responses to the questionnaire for the entire study sample. The boxes coloured in red or yellow represent the distribution of the variables describing the probability of manifesting a certain behaviour; boxes in green display the variables describing preferences for certain characteristics of the route. Variables coloured in yellow were included in the clustering analysis. *UE*: urban elements; *GMH*: global minimisation heuristic; *LMH*: local minimisation heuristic. (For interpretation of the references to colour in this figure caption, the reader is referred to the Web version of this article)

Table 3
The Silhouette scores and VRC coefficients ω of the clustering structures generated by the k-means algorithm, with different desired numbers of clusters as input.

Nr. clusters	Silhouette score	coefficient ω
3	0.39	−127.48
4	0.36	120.29
5	0.38	43.51
6	0.39	−22.78
7	0.39	−14.31
8	0.36	−19.14
9	0.36	36.85

influenced by distant landmarks (median probability 0.36), regions (0.26), barriers (0.25), and, to a lesser extent, on-route marks (0.21).[3]

Looking at the minimisation heuristics, one can observe a preference for road distance minimisation over cumulative angular change minimisation, both at the global (median probabilities: 0.22 *vs* 0.13) and local level (0.33 *vs* 0.28). This result is consistent with previous research according to which urban explorers at times minimise road costs (e.g. Gärling et al., 1986; Guo & Loo, 2013; Rodríguez et al., 2015); at the same time, such findings question the great attention (e.g. Esposito et al., 2020; Jiang & Jia, 2011; Omer & Kaplan, 2017) that route choice models based on the least cumulative angular change have received in the design of pedestrian urban agents.

Finally, in Section V, people reported that walking in the proximity of natural elements, on the one hand, and avoiding severing barriers and unfriendly paths along vehicular traffic, on the other hand, are crucial factors in determining where they walk (median of 0.50 for both variables, on the scale 'not important' (0.0), 'slightly important', 'important', 'fairly important', 'very important' (1.0)). However, aversion to severing barriers presents a higher variation (standard deviation: 0.30) than the preference for natural barriers (0.19); this indicates a more widely shared appreciation for natural barriers in contrast to the negative properties associated with the role of severing barriers in hindering walking behaviour.

3.2. Cluster analysis

We found that the silhouette scores obtained from the iterations between 3 and 9 desired clusters ranged from 0.36 to 0.39 (Table 3); this may indicate the emergence of structures that are not particularly strong in terms of differences between clusters (Struyf et al., 1996). Nevertheless, the scores indicate the existence of a division that justifies further investigation. In light of the coefficient ω obtained from the VRC scores and a subjective examination of the clusters, we adopted the partition resulting from k-means with 6 desired clusters. This partition is the second-best in terms of coefficient ω, it presents more substantial differences between groups, and the sizes of its clusters are reasonable, i.e. no clusters with very few individuals ($N < 10$) were generated.

The chosen structure presents three large clusters composed of 91, 60, and 67 individuals (clusters 2, 3, and 4), and three less populated clusters (clusters 1, 5, and 6, with 31, 21, and 31 members, respectively) (Fig. 5 and Table 4). Clusters 2, 3, and, to a lower extent, 4, are mainly shaped by the attributes at the bottom of the plot: the employment of distant landmarks and barrier sub-goals. *Cluster 2* relies on most of the elements that structure the Image of the City from a global frame of reference: distant landmarks (0.40), environmental barriers (0.32), and, at times, regions (0.28); its members also minimise road distance at the local level. Subjects in *cluster 3* take advantage of distant landmarks (0.45) and, more rarely, barriers (0.26); this cluster includes the most "utilitarian" subjects of the sample (probability of minimising angular change or road distance globally equal to 0.41). Finally, *cluster 4* is featured by a large usage of distant landmarks (0.35) and regions (0.34).

Conversely, the attributes of clusters 1, 5, and 6 depict a diagonal shape towards the top of the plot (high probability of using on-route marks). The subjects belonging to *cluster 1* are likely to walk between on-route marks (0.35) or barrier sub-goals (0.26) by locally minimising road distance. Yet, individuals in this group rarely minimise road costs globally. Subjects in *cluster 5* tend to segment the route on the basis of on-route marks (0.43) and, occasionally, regional divisions (0.32); they do not disdain minimising distance at the global level. The shape of *cluster 6* extends itself from the top to the button of the plot, indicating a concurrent usage of on-route marks (0.36) and distant landmarks (0.28), along with global cumulative angular change minimisation; thus, cluster 6 can be seen as a landmark-based group.

Overall, as concerns the interplay amongst urban elements, two behavioural spectra can be distinguished. On the one hand, there is a set of clusters (clusters 2, 3, and 4) wherein individuals rely on the usage of elements that shape the cognitive map of the city at the global level; on the other hand, there is a portion of individuals, allocated to clusters 1, 5, and 6, who tend to segment their routes by using sub-goals (mostly local landmarks). In the first case, substantial distant landmark

[3] The UE probabilities overlap and thus they represent the concurrent usage of more than one urban element.

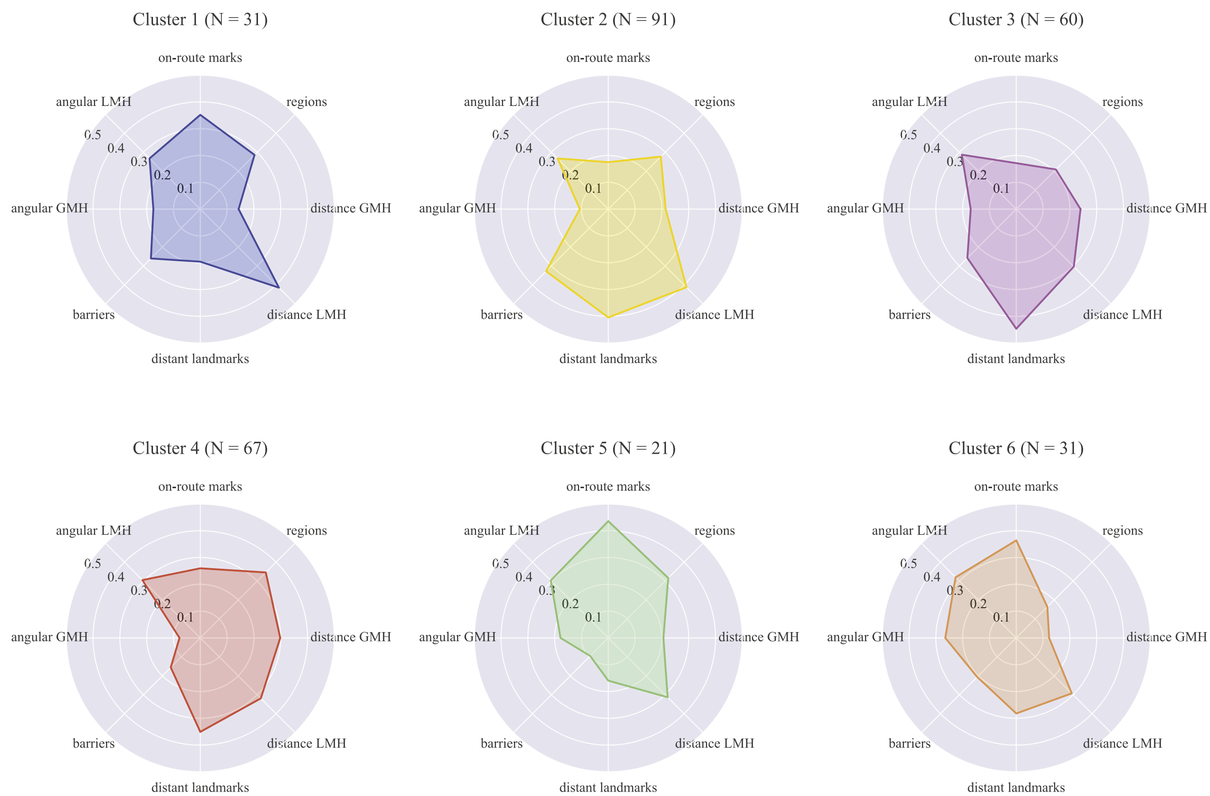

Fig. 5. The clusters resulting from the chosen structure (k-means method, 6 clusters) and the mean values of the variable considered.

Table 4
Demographic composition of the clusters and mean values of the variables not employed in the cluster analysis. Self-reported spatial orientation skills were also summarised by averaging the responses to Question 4 in Section I of the questionnaire.

Cluster	Demographic composition				Preference natural barriers	Aversion to severing barriers	Spatial skills
	Age	Female	Male	Other			
1	28.03	64.5%	35.5%	–	0.46	0.52	3.58 (SD: 0.85)
2	28.17	56.0%	44.0%	3%	0.50	0.52	3.90 (SD: 0.87)
3	27.22	72.5%	26.5%	1.0%	0.50	0.46	3.77 (SD: 0.74)
4	28.64	56.7%	41.8%	1.5%	0.50	0.64	3.97 (SD: 0.74)
5	25.43	71.4%	28.6%	–	0.43	0.50	3.40 (SD: 0.82)
6	27.93	67.7%	29.0%	3.3%	0.47	0.56	3.52 (SD: 0.94)

probabilities are paired with a moderate likelihood of using barriers (cluster 2) or regions (clusters 3 and 4) rather than on-route marks. This could indicate that, at least to describe the behaviour of a large fraction of pedestrians, landmark-based piloting should be broken up into on-route marks and distant landmark-based strategies, instead of being considered a route choice mechanism in itself (see Epstein & Vass, 2014). Such a result may be interpreted by considering the fundamental role of barriers and distant landmarks in structuring people's cognitive map and the city's imageability (Cadwallader, 1976). A correlation[4] of 0.52 between the probability of using distant landmarks and barrier sub-goals over the entire study sample further supports this supposition. Furthermore, the clusters that rely on distant landmarks are also the ones in which subjects are more likely to consider a segmentation based on the perception of regions, another component of people's cognitive maps that contributes to a bird's-eye view of the city (as maintained in Allen & Golledge, 2007; Golledge, 1999).

On the contrary, clusters 1, 5, and 6 are composed of subjects who need path-maintaining local clues (Siegel & White, 1975) to shape their route. These groups present high probabilities of relying on on-route marks and minimising road distance locally. They also exhibit the lowest probabilities of minimising road distance globally. One could argue that individuals in such clusters do not take on a global perspective on the city, that is, they do not embrace an allocentric frame of reference. Instead, they use the information and elements available along the route. The participants in clusters 2, 3, and 4 are more capable of minimising distance globally, whereas the participants belonging to clusters 1, 5, and 6 need to identify and establish relationships between sub-goals to use metric information. The firsts seem to adopt a navigation approach based on the *usage of cognitive maps*, in contrast to *route-based navigation*, that better integrates information about regional divisions and relationships between locations (see Allen, 1999; Allen &

[4] Pearson product-moment correlation coefficient.

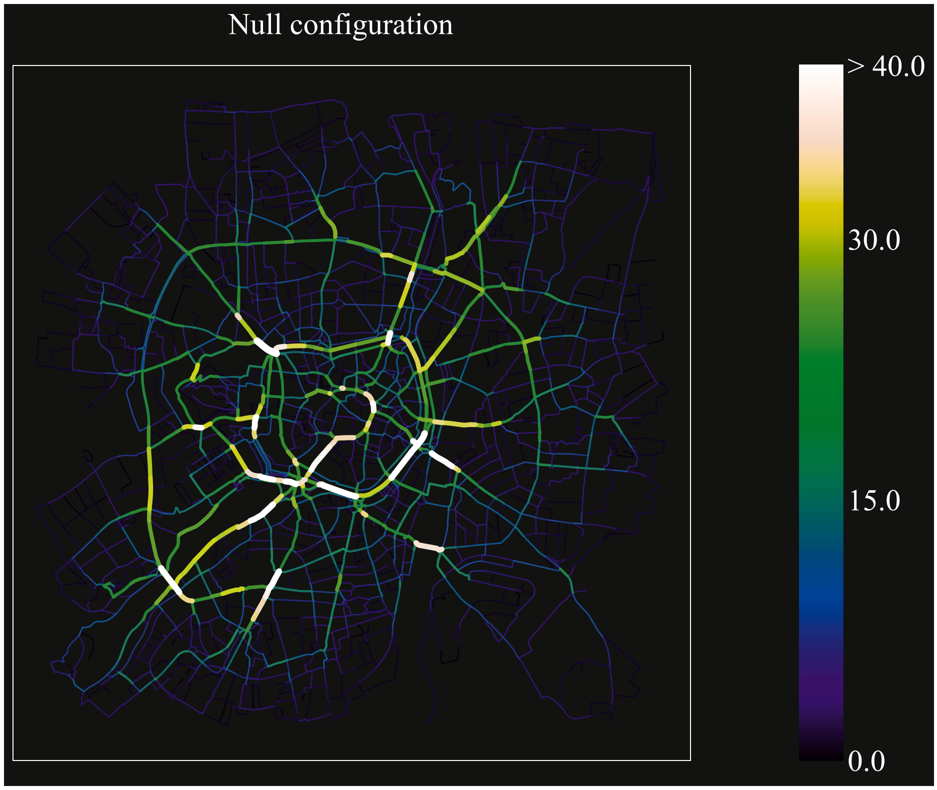

Fig. 6. Movement flows of pedestrian agents across the street network resulting from the null configuration (volumes per street segment, median across runs). (For interpretation of the references to colour in this figure colour bar, the reader is referred to the Web version of this article)

Table 5
Statistics of the ABM configurations.

	Null conf.	Homogeneous conf.	Heterogeneous conf.
Mean volume per street	6.58	6.75	6.74
SD volume per street	7.93	8.16	8.13
Max volume per street	49.50	61.00	62.00
Median route length	2477.92 m.	2523.55 m.	2527.17 m

Golledge, 2007; Epstein & Vass, 2014, for definitions).[5] Cluster 6 manifests a peculiar situation wherein both distant and on-route marks are employed, hence resembling the mechanisms of a complete landmark-based piloting approach (Epstein & Vass, 2014), modelled by Filomena and Verstegen (2021).

In terms of road costs, while clusters 1 and 2 are the groups that minimise road costs the least, cluster 3 and 6 are the most utilitarian. Overall, the probabilities of following the routes generated exclusively by the least cumulative angular change algorithm are lower than 0.20, except for cluster 6 (0.26). This route choice strategy does not appear to be a strong determinant of movement within the study sample, at least for medium and long walking trips. This evidence aligns with the results obtained for a larger case study area (London (UK); see Filomena & Verstegen, 2021). However, the more balanced relationship between cumulative angular change and road distance when employed as local heuristics is consistent with previous findings showing a good prediction capability of the least cumulative angular change algorithm for short pedestrian trips (Omer & Kaplan, 2017).

3.3. The pedestrian flows

The null configuration (Fig. 6 and Table 5) generated pedestrian movement flows that, considering our direct knowledge of the case study area and existing research on pedestrian preferences for certain properties of the environment (e.g. Forsyth et al., 2008; Owen et al., 2004; Sarkar et al., 2015), appear to be already plausible: the lake (location 4 in Fig. 3), the main street (Prinzipalmarkt) in the historical centre, and the Promenade (location 8) - a pedestrian and bike avenue extending around the historical centre - present high volumes of pedestrians (between 40 and 50 trips, around 5% of the total number of trips). Other roads featured by a multitude of amenities, restaurants, and a lively atmosphere were also often traversed (e.g. Wolbecker Straße, location 3, Warendorfer Straße, location 2, and Hammer Straße, location

[5] This does not indicate that individuals exhibiting a route-based navigation approach do not use their cognitive maps. Rather, navigation based on the usage of cognitive maps refers to a subject's ability to take on an allocentric frame of reference and a more structured cognitive representation of space.

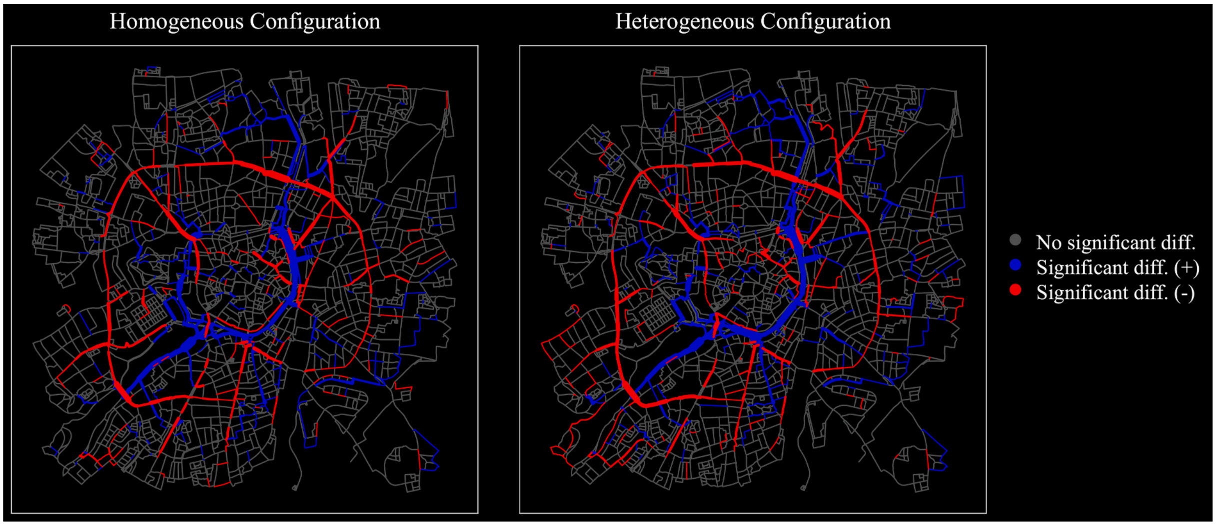

Fig. 7. Statistically significant differences from the null configuration in the movement flows of pedestrian agents across the street network, homogeneous and heterogeneous configurations (pedestrian volumes per street segment, median across runs); "+" indicates segments for which the volumes generated by the homogeneous or heterogeneous configuration are higher than the null configuration; "-" indicates segments with lower volumes. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

10, 15 to 40 trips). The prominence of these roads is also due to their role in connecting outer districts with the inner city. Not only are some of them links between different regions, but others are also leading to or partly extending along natural barriers. At the same time, in the null configuration, many agents employed the outer ring (location 1), a thoroughfare designed for cars and featured by a multilane street design. This is evident in the south-west of the city, near the lake, and towards the north-east (20–35 trips).

Such a behaviour can be, on the one hand, the result of agents minimising cumulative angular change (the ring's shape and its relationship with the rest of the network are favourable characteristics in inducing smooth changes of direction); on the other hand, in the north-east in particular, the ring represents one of the few options to walk between different districts, either because of the separating presence of the railway or because of a not particularly dense system of alternative minor roads. As the ring is a prominent structuring barrier, and does not entail abrupt changes of direction, it may represent the simplest option from a cognitive perspective. However, excluding the volumes in the vicinity of the lake, the rest of the pedestrian volumes appear excessively high for such an unfriendly pedestrian thoroughfare.

The volumes of both the homogeneous and heterogeneous configurations (Fig. 7; see also Fig. A1 in the Appendix for the pedestrian volumes emerging from each agent typology) are significantly lower along the entire extension of the ring; the highest differences from the null configuration (-11 and -10 trips, respectively) are observable in the south-west (location 12) and north-east (location 1) of the ring. Moreover, the empirical configurations brought up significantly lower volumes along Weseler Straße (location 6), Steinfurter Straße (location 11), and the segments that branch out from it towards the ring; this is a set of highly congested roads, not particularly friendly to pedestrians.

Concurrently, as opposed to the null configuration, the homogeneous and heterogeneous configurations feature a significantly higher number of trips along the east bank of the lake (location 4, $+13$ and $+ 14$ trips) and in the area between the lake and the Promenade (location 9, $+6$ and $+ 11$). Along the entire length of the Promenade, the empirical configurations present significantly higher volumes of pedestrians (locations 8, $+13$ trips), apart from the north-west section. It is worth mentioning that in the east ($+11$) and south-east segments of the Promenade (location 7, $+12$) the heterogeneous configuration displays more remarkable differences. Finally, the empirically based configurations resulted in significantly larger volumes in other locations such as the west and east sides of the university's historical castle (location 5), and slightly lower volumes in a few roads in the historical centre.

The paths along the Promenade and the lake, because of their characteristics - surrounded by green and water elements, and in the proximity of historical landmarks - are in reality very attractive to pedestrians in Münster, both locals and tourists who may be looking for a walk along restorative natural elements. As the volumes of the null configuration would not draw enough attention to the prominence of such paths in terms of pedestrian distributions, we claim that the patterns emerging from the empirically based configurations are more plausible and informative for decision-makers and urban planners. This is also emphasised by the fact that most of the street segments that feature significantly lower volumes in the empirically based configurations are in fact traversed by considerable volumes of vehicular traffic (major and secondary roads; see also Fig. 3). In light of the values of the stochastic preference parameters,[6] one can argue that the divergence between the volumes of the empirically based and the null configurations derived from how the values of the stochastic discrete parameters contributed to defining the agents' route choice behaviour and not from the barrier effect on the perception of road costs.

Although we identified crucial areas where the empirical configurations produced significant and meaningful differences in pedestrian movement patterns compared to a non-empirically based ABM, the heterogeneous configuration does not substantially differ from the homogeneous configuration, besides a few negligible differences. Multiple runs of the model might have flattened out the differences, especially on the global urban scale; this phenomenon has been partially observed in previous work on residential choice (see Brown & Robinson, 2006; Buchmann et al., 2016). Thus, a major effort is required in simulation models of pedestrian movement in cities to identify key individual aspects for which the inclusion of heterogeneity may affect pedestrian flow patterns. For example, although on a different scale of analysis, Haghani and Sarvi (2017) found that accounting for heterogeneity in the agent's

[6] On average, segments adjacent to severing barriers were perceived as 50% longer in the null configuration and 54% longer in the empirically based configurations; segments along natural barriers were perceived as 50% shorter in all configurations.

perception of peer influence led to low prediction errors in a crowd simulation model of escape behaviour. Heterogeneity may tend to become more influential on the outputs of the model when the agents interact with each other, hence following feedback and path dependence mechanisms, that is, how variation in initial conditions or interactions between the ABM components may lead to remarkable deviations in the resulting patterns (Brown et al., 2005; Railsback & Grimm, 2019). Agents belonging to the same typology can be designed to be influenced by how "peers", identified on the basis of similar goals and representations, act. These dynamics can allow analysing the effects of interactions between different groups across the pedestrian population (e.g. tourists *vs* students *vs* commuters, etc.).

3.4. Limitations and future work

The research discussed in this paper presents limitations as concerns the data collection method and sample, and the validation of the pedestrian volumes emerging from the ABM configuration. Instead of the core tasks of the questionnaire, whereby the participants were asked to formulate routes through video sequences, one could design a more ecological experiment by asking the subjects to walk between two locations. While this approach would grant more freedom to the participants as regards their choices, it would also be more time consuming, both in terms of participant recruitment and actual execution. More importantly, it would make the procedure of linking the results to the parameter settings of the model less straightforward. Furthermore, the study sample mainly comprises young adults, either in possession of a degree or pursuing one. As such, it is not representative of the different age groups in the target population (people in the 15–29 years age group correspond to 25% of the population in Münster; (see StatistischeÄmter des Bundes und der Länder, 2021), neither it accounts for differences in behaviour that may be related to sex (52% female and 48% male in Münster, *vs* 63% and 37% in our study; see Statistische Ämter des Bundes und der Länder, 2021),[7] educational level, nor cultural differences.

The evaluation of an ABM would benefit from a validation of the emerging patterns. Pedestrian volume data should be employed to verify that the model can generate realistic patterns. However, acquiring precise pedestrian counts for a sufficient number of road segments would require deploying a large set of humans (manual counts) or resources for setting up automatic counting devices. City administrations that have spread such devices across their roads are rare; instead, pedestrian movement data, such as trajectories or mobile phone location data, are in the hands of private companies, unwilling to share them.

Finally, agents should be endowed with social interaction capabilities; amongst others, this aspect defines truly cognitive agents (Castelfranchi, 2000; Conte & Paolucci, 2014) and distinguishes agent-based simulations from other mathematical models (Epstein, 1999). Moreover, it would give life to more complex and interactive path-dependence mechanisms, as described above. In this and previous work, we have focused on devising the representational capabilities of pedestrian agents; we believe that the inclusion of social situatedness and social intelligence should motivate future research in pedestrian movement simulation.

4. Conclusion

The aim of this paper was to introduce an empirically based agent-based model (Rounsevell et al., 2012) for the simulation of pedestrian movement in urban areas that contemplates behavioural heterogeneity in pedestrian route choice strategies. The following research questions drove our investigation: a) What is the diversity in route choice strategies as concerns the usage of minimisation heuristics and information about meaningful urban elements? b) To what extent does the variation in the agents' heterogeneity - i.e. a model including agent typologies *vs* a model with a homogeneous set of agents - generate different movement patterns across the street network? The city of Münster (DE) was used as a case study area.

To address the first research question, we investigated with a questionnaire how pedestrians may differ in the adoption of route choice strategies in relation to the use of salient urban elements (landmarks, regions, and barriers) and road costs. We identified six clusters from a sample composed of 301 subjects. Three clusters showed behavioural patterns associated with the usage of global structuring elements (distant landmarks, urban subdivisions); members of the other three clusters, in contrast, resorted to local elements identified along the route (on-route marks, barrier sub-goals at times).

To answer the second research question, three different ABM configurations were devised: a null configuration, wherein the agent behaviour was regulated by randomly extracted values; a homogeneous configuration, in which the study sample attributes obtained from the questionnaire were employed to direct the behaviour of the agents; a heterogeneous configuration, whereby agent typologies were built to diversify the agent behaviour. The ABM configurations generated pedestrian movement flows across the street network that we deem plausible. Most of the agents concentrated near natural elements and traversed the streets of the historical centre or other lively roads. However, as opposed to the null configuration, the empirically based configurations showed significantly lower volumes along major thoroughfares designed for cars and produced higher volumes in areas that, in reality, present substantial pedestrian concentrations. Very few differences were identified instead between the pedestrian flows emerging from the homogeneous and heterogeneous configurations.

This work represents the first effort to devise an empirically based ABM for the simulation of pedestrian movement in cities. It provides a comprehensive approach to inform and ground the conception of cognitive maps for pedestrian agents. Not only have we shown that individuals make large use of a diversified set of urban elements - providing further evidence on the limits of rational and utilitarian approaches to human cognition - but also that the interaction between the urban environment and people's route choice strategies is not static and homogeneous, but, rather, dynamic (an individual may use different elements over different trips) and multifaceted.

[7] Estimates based on the 2011's census data.

Appendix

A.1. Data availability

A Python Jupyter Notebook presenting the cluster analysis and a second Jupyter Notebook documenting the evaluation of the ABM outcomes are available on a Github repository. Therein, the original questionnaire and its response table are also contained (Filomena, 2022a). Along with the case study input data, the agent-based model is freely available on a GitHub repository (Filomena, 2022b).

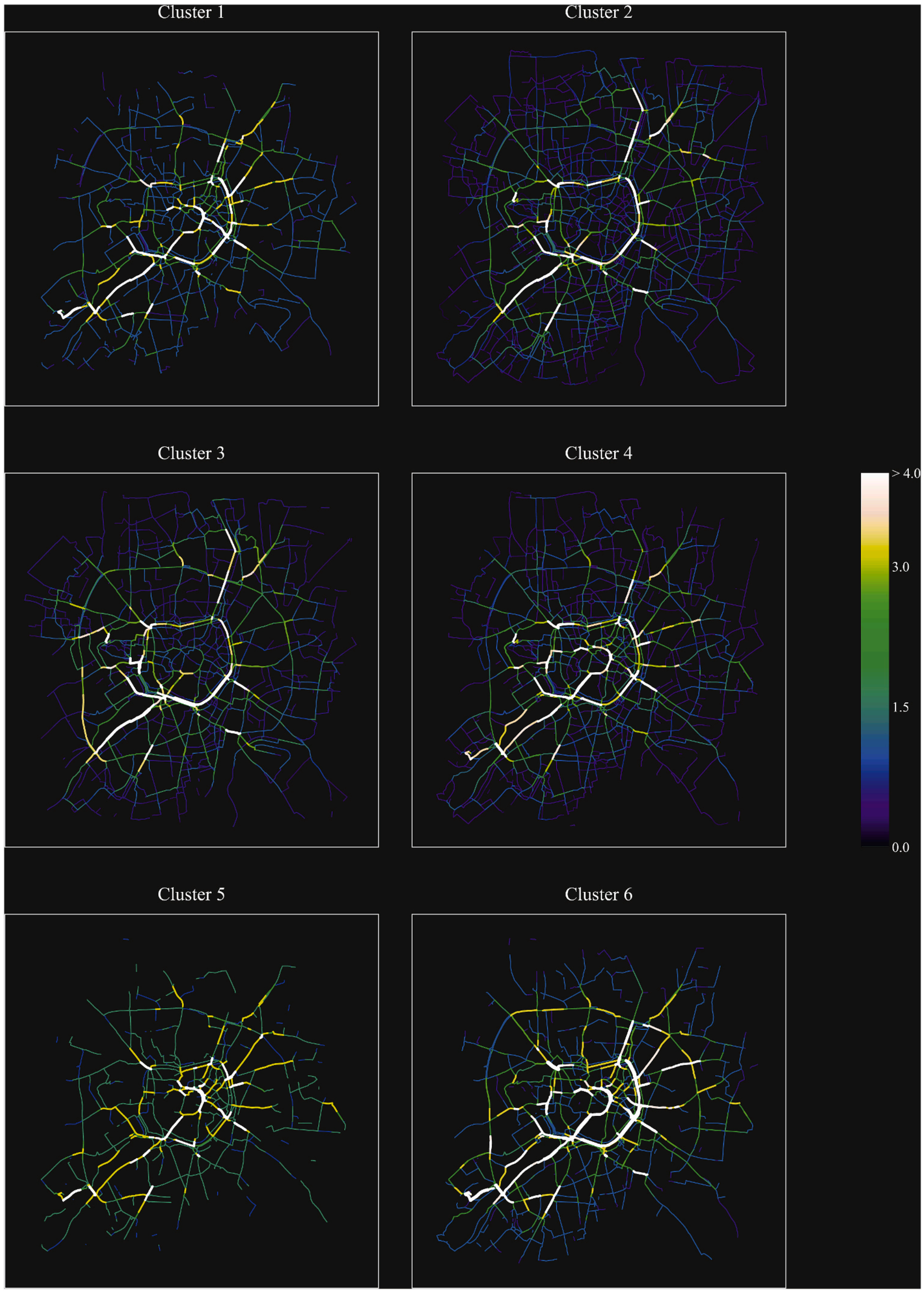

Fig. A1. Movement flows of pedestrian agents across the street network resulting from the volumes of each agent typology in the heterogeneous configuration (percentage of agent volumes per street segment, median across runs). (For interpretation of the references to colour in this figure colour bar, the reader is referred to the Web version of this article).

References

Alfonzo, M. A. (2005). To walk or not to walk? The hierarchy of walking needs. *Environment and Behavior, 37*, 808–836. https://doi.org/10.1177/0013916504274016

Allen, G. L. (1999). Spatial abilities, cognitive maps, and wayfinding - Bases for individual differences in spatial cognition and behavior. In R. G. Golledge (Ed.), *Wayfinding behavior: Cognitive mapping and other spatial processes* (pp. 46–80). Baltimore, MD: Johns Hopkins University Press.

Allen, G. L., & Golledge, R. G. (2007). Wayfinding in Urban environments. In T. Gärling, & L. Steg (Eds.), *Threats from car traffic to the quality of urban life* (pp. 79–101). Amsterdam: Elsevier. https://doi.org/10.1108/9780080481449.

Bonabeau, E. (2002). Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences, 99*, 7280–7287. https://doi.org/10.1073/pnas.082080899

Bornioli, A., Parkhurst, G., & Morgan, P. L. (2018). Psychological wellbeing benefits of simulated exposure to five Urban settings: An experimental study from the pedestrian's perspective. *Journal of Transport & Health, 9*, 105–116. https://doi.org/10.1016/j.jth.2018.02.003

Brown, D. G., Page, S., Riolo, R., Zellner, M., & Rand, W. (2005). Path dependence and the validation of agent-based spatial models of land use. *International Journal of Geographical Information Science, 19*, 153–174. https://doi.org/10.1080/13658810410001713399

Brown, D. G., & Robinson, D. T. (2006). Effects of heterogeneity in residential preferences on an agent-based model of urban sprawl. *Ecology and Society, 11*. https://doi.org/10.5751/ES-01749-110146

Buchmann, C. M., Grossmann, K., & Schwarz, N. (2016). How agent heterogeneity, model structure and input data determine the performance of an empirical ABM - A real-world case study on residential mobility. *Environmental Modelling & Software, 75*, 77–93. https://doi.org/10.1016/j.envsoft.2015.10.005

Cadwallader, M. (1976). Cognitive distance in intraurban space. In G. T. Moore, & R. G. Golledge (Eds.), *Environmental knowing* (pp. 316–324). Stroudsburg, PA: Dowden, Hutchinson & Ross.

Caliński, T., & Harabasz, J. (1974). A dendrite method for cluster Analysis. *Communications in Statistics, 3*, 1–27.

Castelfranchi, C. (2000). Through the agents' minds: Cognitive mediators of social action. *Mind & Society, 1*, 109–140. https://doi.org/10.1007/BF02512232

Conte, R., & Paolucci, M. (2014). On agent-based modeling and computational social science. *Frontiers in Psychology, 5*, 1–9. https://doi.org/10.3389/fpsyg.2014.00668

Crooks, A. T., Castle, C., & Batty, M. (2008). Key challenges in agent-based modelling for geo-spatial simulation. *Computers, Environment and Urban Systems, 32*, 417–430. https://doi.org/10.1016/j.compenvurbsys.2008.09.004

Darker, C. D., Larkin, M., & French, D. P. (2007). An exploration of walking behaviour - An interpretative phenomenological approach. *Social Science & Medicine, 65*, 2172–2183. https://doi.org/10.1016/j.socscimed.2007.06.029

De Certeau, M. (1984). *The practice of everyday life*. Berkeley, CA: University of California Press.

Dovey, K., & Pafka, E. (2014). The urban density assemblage: Modelling multiple measures. *Urban Design International, 19*, 66–76. https://doi.org/10.1057/udi.2013.13

Duives, D. C., Daamen, W., & Hoogendoorn, S. P. (2013). State-of-the-art crowd motion simulation models. *Transportation Research Part C: Emerging Technologies, 37*, 193–209. https://doi.org/10.1016/J.TRC.2013.02.005

Epstein, J. M. (1999). Agent-based computational models and generative social science. *Complexity, 4*, 41–60. https://doi.org/10.1002/(SICI)1099-0526(199905/06)4:5<41::AID-CPLX9>3.0.CO;2-F

Epstein, R. A., & Vass, L. K. (2014). Neural systems for landmark-based wayfinding in humans. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369*. https://doi.org/10.1098/rstb.2012.0533

Esposito, D., Santoro, S., & Camarda, D. (2020). Agent-based analysis of Urban spaces using Space Syntax and spatial cognition approaches: A case study in Bari, Italy. *Sustainability, 12*, 4625. https://doi.org/10.3390/su12114625

Evans, J., & Jones, P. (2011). The walking interview: Methodology, mobility and place. *Applied Geography, 31*, 849–858. https://doi.org/10.1016/j.apgeog.2010.09.005

Everitt, B. S., Landau, S., Leese, M., & Stahl, D. (2011). *Cluster Analysis* (5th ed.). Chichester, UK: John Wiley & Sons.

Ferdman, A. (2019). Walking and its contribution to objective well-being. *Journal of Planning Education and Research*. https://doi.org/10.1177/0739456X19875195, 0739456X1987519.

Fernandez, L. E., Brown, D. G., Marans, R. W., & Nassauer, J. I. (2005). Characterizing location preferences in an exurban population: Implications for agent-based modeling. *Environment and Planning B: Planning and Design, 32*, 799–820. https://doi.org/10.1068/b3071

Filomena, G. (2022a). g-filomena/PedSimCity-Evaluation: PedSimCity-Empirical Based Agent-Based Model: Preparation and Evaluation (v1.05). *Zenodo*. https://doi.org/10.5281/zenodo.6574818. https://github.com/g-filomena/PedSimCity-Evaluation/tree/v1.05

Filomena, G. (2022b). g-filomena/pedsimcity: PedSimCity-Empirical based Agent-Based Model (v1.11). *Zenodo*. https://doi.org/10.5281/zenodo.6574825. https://github.com/g-filomena/pedsimcity/tree/v1.11

Filomena, G., Manley, E., & Verstegen, J. A. (2020). Perception of urban subdivisions in pedestrian movement simulation. *PLoS One, 15*, Article e0244099. https://doi.org/10.1371/journal.pone.0244099

Filomena, G., & Verstegen, J. A. (2021). Modelling the effect of landmarks on pedestrian dynamics in urban environments. *Computers, Environment and Urban Systems, 86*, Article 101573. https://doi.org/10.1016/j.compenvurbsys.2020.101573

Filomena, G., Verstegen, J. A., & Manley, E. (2019). A computational approach to 'The Image of the City. *Cities, 89*, 14–25. https://doi.org/10.1016/j.cities.2019.01.006

Foltête, J. C., & Piombini, A. (2010). Deviations in pedestrian itineraries in urban areas: A method to assess the role of environmental factors. *Environment and Planning B: Planning and Design, 37*, 723–739. https://doi.org/10.1068/b35015

Formann, A. K. (1984). *Die Latent-Class-Analyse: Einführung in die Theorie und Anwendung [Latent class analysis: Introduction to theory and application]*. Weinheim: Beltz.

Forsyth, A., Hearst, M., Oakes, J. M., & Schmitz, K. H. (2008). Design and destinations: Factors influencing walking and total physical activity. *Urban Studies, 45*, 1973–1996. https://doi.org/10.1177/0042098008093386

Franklin, S., & Graesser, A. (1996). Is it an agent, or just a program? A taxonomy of autonomous agents. In J. Müller, M. Wooldridge, & N. Jennings (Eds.), *Lecture Notes in Computer Science: vol. 1193. Intelligent Agents III Agent Theories, Architectures, and Languages. ATAL 1996* (pp. 21–35). Berlin, Heidelberg: Springer.

Gärling, T., Säisä, J., Book, A., & Lindberg, E. (1986). The spatiotemporal sequencing of everyday activities in the large-scale environment. *Journal of Environmental Psychology, 6*, 261–280. https://doi.org/10.1016/S0272-4944(86)80001-9

Gatersleben, B., & Andrews, M. (2013). When walking in nature is not restorative - The role of prospect and refuge. *Health & Place, 20*, 91–101. https://doi.org/10.1016/j.healthplace.2013.01.001

Gidlow, C. J., Jones, M. V., Hurst, G., Masterson, D., Clark-Carter, D., Tarvainen, M. P., Smith, G., & Nieuwenhuijsen, M. (2016). Where to put your best foot forward: Psycho-physiological responses to walking in natural and urban environments. *Journal of Environmental Psychology, 45*, 22–29. https://doi.org/10.1016/j.jenvp.2015.11.003

Golledge, R. G. (1995). Path selection and route preference in human navigation: A progress report. In A. U. Frank, & W. Kuhn (Eds.), *Lecture Notes in Computer Science: vol. 988. Spatial information theory A theoretical basis for GIS. COSIT 1995* (pp. 207–222). Berlin, Heidelberg: Springer. https://doi.org/10.1007/3-540-60392-1_14.

Golledge, R. G. (1999). Human wayfinding and cognitive maps. In R. G. Golledge (Ed.), *Wayfinding behavior: Cognitive mapping and other spatial processes* (pp. 5–45). Baltimore, MD: Johns Hopkins University Press.

Guo, Z., & Loo, B. P. (2013). Pedestrian environment and route choice: Evidence from New York city and Hong Kong. *Journal of Transport Geography, 28*, 124–136. https://doi.org/10.1016/j.jtrangeo.2012.11.013

Haacke, H. C., Enßle, F., Haase, D., & Lakes, T. (2022). *How to derive spatial agents: A mixed-method approach to model an elderly population with scarce data. Population, Space and Place* (p. e2551). https://doi.org/10.1002/psp.2551

Haghani, M. (2020a). Empirical methods in pedestrian, crowd and evacuation dynamics: Part I. Experimental methods and emerging topics. *Safety Science, 129*, Article 104743. https://doi.org/10.1016/J.SSCI.2020.104743

Haghani, M. (2020b). Empirical methods in pedestrian, crowd and evacuation dynamics: Part II. Field methods and controversial topics. *Safety Science, 129*, Article 104760. https://doi.org/10.1016/j.ssci.2020.104760

Haghani, M., & Sarvi, M. (2017). How perception of peer behaviour influences escape decision making: The role of individual differences. *Journal of Environmental Psychology, 51*, 141–157. https://doi.org/10.1016/j.jenvp.2017.03.013

Handy, S. L. (1996). Urban form and pedestrian choices: Study of Austin neighborhoods. *Transportation Research Record: Journal of the Transportation Research Board, 1552*, 135–144. https://doi.org/10.1177/0361198196155200119

Han, J., Kamber, M., & Pei, J. (2012). Cluster Analysis: Basic concepts and methods. In *Data mining concepts and techniques* (3rd ed., pp. 443–495). Waltham, MA: Morgan Kaufmann Publishers. https://doi.org/10.1016/B978-0-12-381479-1.00010-1.

Heppenstall, A., Malleson, N., & Crooks, A. T. (2016). Space, the final frontier": How good are agent-based models at simulating individuals and space in cities? *Systems, 4*, 9. https://doi.org/10.3390/systems4010009

Hill, M. R. (1984). *Walking, crossing streets and choosing pedestrian routes: A survey of recent insights from the social/Behavioral sciences*. Lincoln, NE: University of Nebraska.

Hölscher, C., Tenbrink, T., & Wiener, J. M. (2011). Would you follow your own route description? Cognitive strategies in urban route planning. *Cognition, 121*, 228–247. https://doi.org/10.1016/j.cognition.2011.06.005

Jager, W. (2017). Enhancing the realism of simulation (EROS): On implementing and developing psychological theory in social simulation. *The Journal of Artificial Societies and Social Simulation, 20*, 14. https://doi.org/10.18564/jasss.3522

Jager, W., & Ernst, A. (2017). Introduction of the special issue "Social simulation in environmental psychology". *Journal of Environmental Psychology, 52*, 114–118. https://doi.org/10.1016/j.jenvp.2017.07.002

Jan, O., Horowitz, A. J., & Peng, Z. R. (2000). Using global positioning system data to understand variations in path choice. *Transportation Research Record*, 37–44. https://doi.org/10.3141/1725-06

Janssen, M. A., & Ostrom, E. (2006). Empirically based, agent-based models. *Ecology and Society, 11*, 37. https://doi.org/10.5751/ES-01861-110237

Jiang, B., & Jia, T. (2011). Agent-based simulation of human movement shaped by the underlying street structure. *International Journal of Geographical Information Science, 25*, 51–64. https://doi.org/10.1080/13658811003712864

Kato, Y., & Takeuchi, Y. (2003). Individual differences in wayfinding strategies. *Journal of Environmental Psychology, 23*, 171–188. https://doi.org/10.1016/S0272-4944(03)00011-2

Kerridge, J., Hine, J., & Wigan, M. (2001). Agent-based modelling of pedestrian movements: The questions that need to Be asked and answered. *Environment and Planning B: Planning and Design, 28*, 327–341. https://doi.org/10.1068/b2696

Kim, H. (2015). Walking distance, route choice, and activities while walking: A record of following pedestrians from transit stations in the San Francisco Bay area. *Urban Design International, 20*, 144–157. https://doi.org/10.1057/udi.2015.2

Koh, P., & Wong, Y. (2013). Influence of infrastructural compatibility factors on walking and cycling route choices. *Journal of Environmental Psychology, 36*, 202–213. https://doi.org/10.1016/j.jenvp.2013.08.001

Lloyd, S. P. (1982). Least squares quantization in PCM. *IEEE Transactions on Information Theory, 28*, 129–137.

Lorimer, H. (2010). Walking: New forms and spaces for studies of pedestrianism. In T. Cresswell, & P. Merriman (Eds.), *Geographies of mobilities: Practices, spaces, subjects* (pp. 19–33). Aldershot: Ashgate. https://doi.org/10.4324/9781315584393-6.

Lynch, K. (1960). *The image of the city*. Cambridge, MA: MIT Press.

Mallot, H. A., & Basten, K. (2009). Embodied spatial cognition: Biological and artificial systems. *Image and Vision Computing, 27*, 1658–1670. https://doi.org/10.1016/j.imavis.2008.09.001
Middleton, J. (2011). Walking in the city: The geographies of everyday pedestrian practices. *Geography Compass, 5*, 90–105. https://doi.org/10.1111/j.1749-8198.2010.00409.x
Middleton, J. (2018). The socialities of everyday urban walking and the 'right to the city. *Urban Studies, 55*, 296–315. https://doi.org/10.1177/0042098016649325
Milligan, G. W., & Cooper, M. C. (1985). An examination of procedures for determining the number of clusters in a data set. *Psychometrika, 50*, 159–179. https://doi.org/10.1007/BF02294245
Münster, S. (2021a). Germany's cycling capital. https://www.stadt-muenster.de/english/ten-times-muenster-in-brief/germanys-cycling-capital.html. (Accessed 5 May 2021).
Münster, S. (2021b). Münster in figures. https://www.stadt-muenster.de/english/ten-times-muenster-in-brief/muenster-in-figures. (Accessed 5 May 2021).
Münster, S. (2021c). The world's most liveable city. https://www.stadt-muenster.de/english/ten-times-muenster-in-brief/the-worlds-most-liveable-city. (Accessed 5 May 2021).
Münzer, S., Fehringer, B. C., & Kühl, T. (2016). Validation of a 3-factor structure of spatial strategies and relations to possession and usage of navigational aids. *Journal of Environmental Psychology, 47*, 66–78. https://doi.org/10.1016/j.jenvp.2016.04.017
Münster zu Fuß. Fußverkehrs-Check Kreuzviertel. https://muenster-zu-fuss.de/2020/10/11/fussverkehrs-check-kreuzviertel/. (Accessed 8 March 2021).
Muraleetharan, T., & Hagiwara, T. (2007). Overall level of service of Urban walking environment and its influence on pedestrian route choice behavior. *Transportation Research Record: Journal of the Transportation Research Board, 2002*, 7–17. https://doi.org/10.3141/2002-02
Omer, I., & Kaplan, N. (2017). Using Space Syntax and agent-based approaches for modeling pedestrian volume at the urban scale. *Computers, Environment and Urban Systems, 64*, 57–67. https://doi.org/10.1016/J.COMPENVURBSYS.2017.01.007
OpenStreetMap contributors. Planet dump. retrieved from https://planet.osm.org https://www.openstreetmap.org. (Accessed 8 March 2021).
Owen, N., Humpel, N., Leslie, E., Bauman, A., & Sallis, J. F. (2004). Understanding environmental influences on walking. *American Journal of Preventive Medicine, 27*, 67–76. https://doi.org/10.1016/j.amepre.2004.03.006
Paas, F. G. W. C. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology, 84*, 429–434. https://doi.org/10.1037/0022-0663.84.4.429
Papadimitriou, E., Yannis, G., & Golias, J. (2009). A critical assessment of pedestrian behaviour models. *Transportation Research Part F: Traffic Psychology and Behaviour, 12*, 242–255. https://doi.org/10.1016/j.trf.2008.12.004
Pucher, J., & Buehler, R. (2010). Walking and cycling for healthy cities. *Built Environment, 36*, 319–414. https://doi.org/10.2148/benv.36.4.391
Railsback, S. F., & Grimm, V. (2019). *Agent-based and individual-based modelling: A practical introduction* (Second edition). Princeton, NJ: Princeton University Press.
Rodríguez, D. A., Merlin, L., Prato, C. G., Conway, T. L., Cohen, D., Elder, J. P., Evenson, K. R., McKenzie, T. L., Pickrel, J. L., & Veblen-Mortenson, S. (2015). Influence of the built environment on pedestrian route choices of adolescent girls. *Environment and Behavior, 47*, 359–394. https://doi.org/10.1177/0013916513520004
Roe, J., & Aspinall, P. (2011). The restorative benefits of walking in urban and rural settings in adults with good and poor mental health. *Health & Place, 17*, 103–113. https://doi.org/10.1016/j.healthplace.2010.09.003
Rounsevell, M. D., Robinson, D. T., & Murray-Rust, D. (2012). From actors to agents in socio-ecological systems models. *Philosophical Transactions of the Royal Society B: Biological Sciences, 367*, 259–269. https://doi.org/10.1098/rstb.2011.0187
Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics, 20*, 53–65. https://doi.org/10.1016/0377-0427(87)90125-7
Sarkar, C., Webster, C., Pryor, M., Tang, D., Melbourne, S., Zhang, X., & Jianzheng, L. (2015). Exploring associations between urban green, street design and walking: Results from the Greater London boroughs. *Landscape and Urban Planning, 143*, 112–125. https://doi.org/10.1016/j.landurbplan.2015.06.013
Shatu, F., Yigitcanlar, T., & Bunker, J. (2019). Shortest path distance vs. least directional change: Empirical testing of Space Syntax and geographic theories concerning pedestrian route choice behaviour. *Journal of Transport Geography, 74*, 37–52. https://doi.org/10.1016/j.jtrangeo.2018.11.005
Siegel, A. W., & White, S. H. (1975). The development of spatial representations of large-scale environments. *Advances in Child Development and Behavior, 10*, 9–55. https://doi.org/10.1016/S0065-2407(08)60007-5
Spiers, H. J., & Maguire, E. A. (2008). The dynamic nature of cognition during wayfinding. *Journal of Environmental Psychology, 28*, 232–249. https://doi.org/10.1016/j.jenvp.2008.02.006
Statistische Ämter des Bundes und der Länder. (2021). Statistische Ämter des Bundes und der Länder. https://www.regionalstatistik.de/genesis/online/.
Steinley, D. (2004). Standardizing variables in K-means clustering. In D. Banks, F. McMorris, P. Arabie, & W. Gaul (Eds.), *Classification, clustering, and data mining applications* (pp. 53–60). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-17103-1_6.
Stratford, E., Waitt, G., & Harada, T. (2020). Walking city streets: Spatial qualities, spatial justice, and democratising impulses. *Transactions of the Institute of British Geographers, 45*, 123–138. https://doi.org/10.1111/tran.12337
Struyf, A., Hubert, M., & Rousseeuw, P. (1996). Clustering in an object-oriented environment. *Journal of Statistical Software, 1*, 1–30. https://doi.org/10.18637/jss.v001.i04
Sullivan, K., Coletti, M., & Luke, S. (2010). *GeoMason: GeoSpatial support for MASON*. VA: Technical Report Department of Computer Science, George Mason University Fairfax. http://cs.gmu.edu/703-993-1530.
Tight, M., Timms, P., Banister, D., Bowmaker, J., Copas, J., Day, A., Drinkwater, D., Givoni, M., Gühnemann, A., Lawler, M., Macmillen, J., Miles, A., Moore, N., Newton, R., Ngoduy, D., Ormerod, M., O'Sullivan, M., & Watling, D. (2011). Visions for a walking and cycling focussed urban transport system. *Journal of Transport Geography, 19*, 1580–1589. https://doi.org/10.1016/j.jtrangeo.2011.03.011
Torrens, P. M. (2012). Moving agent pedestrians through space and time. *Annals of the Association of American Geographers, 102*, 35–66. https://doi.org/10.1080/00045608.2011.595658
Torrens, P. M. (2016). Computational streetscapes. *Computation, 4*, 1–38. https://doi.org/10.3390/computation4030037
Wiener, J. M., & Mallot, H. A. (2003). 'Fine-to-Coarse' route planning and navigation in regionalized environments. *Spatial Cognition and Computation, 3*, 331–358. https://doi.org/10.1207/s15427633scc0304_5
Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometric Bulletin, 1*, 80. https://doi.org/10.2307/3001968
Willis, A., Gjersoe, N., Havard, C., Kerridge, J., & Kukla, R. (2004). Human movement behaviour in Urban spaces: Implications for the design and modelling of effective pedestrian environments. *Environment and Planning B: Planning and Design, 31*, 805–828. https://doi.org/10.1068/b3060
Yang, Y., & Diez-Roux, A. V. (2012). Walking distance by trip purpose and population subgroups. *American Journal of Preventive Medicine, 43*, 11–19. https://doi.org/10.1016/j.amepre.2012.03.015
Yang, S., Li, T., Gong, X., Peng, B., & Hu, J. (2020). A review on crowd simulation and modeling. *Graphical Models, 111*, Article 101081. https://doi.org/10.1016/j.gmod.2020.101081
Zacharias, J. (2001). Pedestrian behavior and perception in Urban walking environments. *Journal of Planning Literature, 16*, 3–18. https://doi.org/10.1177/08854120122093249
Zhu, S., & Levinson, D. (2015). Do people use the shortest path? An empirical test of Wardrop's first principle. *PLoS One, 10*, Article e0134322. https://doi.org/10.1371/journal.pone.0134322