Delft University of Technology

Unravelling Urban Wayfinding

Studies on the development of spatial knowledge, activity patterns, and route dynamics of cyclists

Zomer, L.

DOI
10.4233/uuid:122bcbc0-72c6-4ab9-8aea-e39716961f46

Publication date
2021

Document Version
Final published version

Citation (APA)
Zomer, L. (2021). *Unravelling Urban Wayfinding: Studies on the development of spatial knowledge, activity patterns, and route dynamics of cyclists.* [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:122bcbc0-72c6-4ab9-8aea-e39716961f46

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

TRAIL

Summary

Every day residents and visitors find their way through the complex urban network to go to work or get education, or go sightseeing. This thesis contains studies on the development of spatial knowledge, activity patterns, and route dynamics of cyclists. The contributions and findings narrowed the gap between research on travel behaviour research and research on urban spatial knowledge.

About the Author

Lara-Britt Zomer conducted her PhD research at Delft University of Technology. She holds a MSc degree in Transport, Infrastructure and Logistics with a specialisation in Operations. Her research interests include unravelling (urban) travel behaviour using mobility data.

TRAIL Research School ISBN 9978-90-5584-291-9

Erasmus University Rotterdam | Radboud University | rijksuniversiteit groningen | TU Delft Delft University of Technology | UNIVERSITY OF TWENTE. | TU/e Technische Universiteit Eindhoven University of Technology

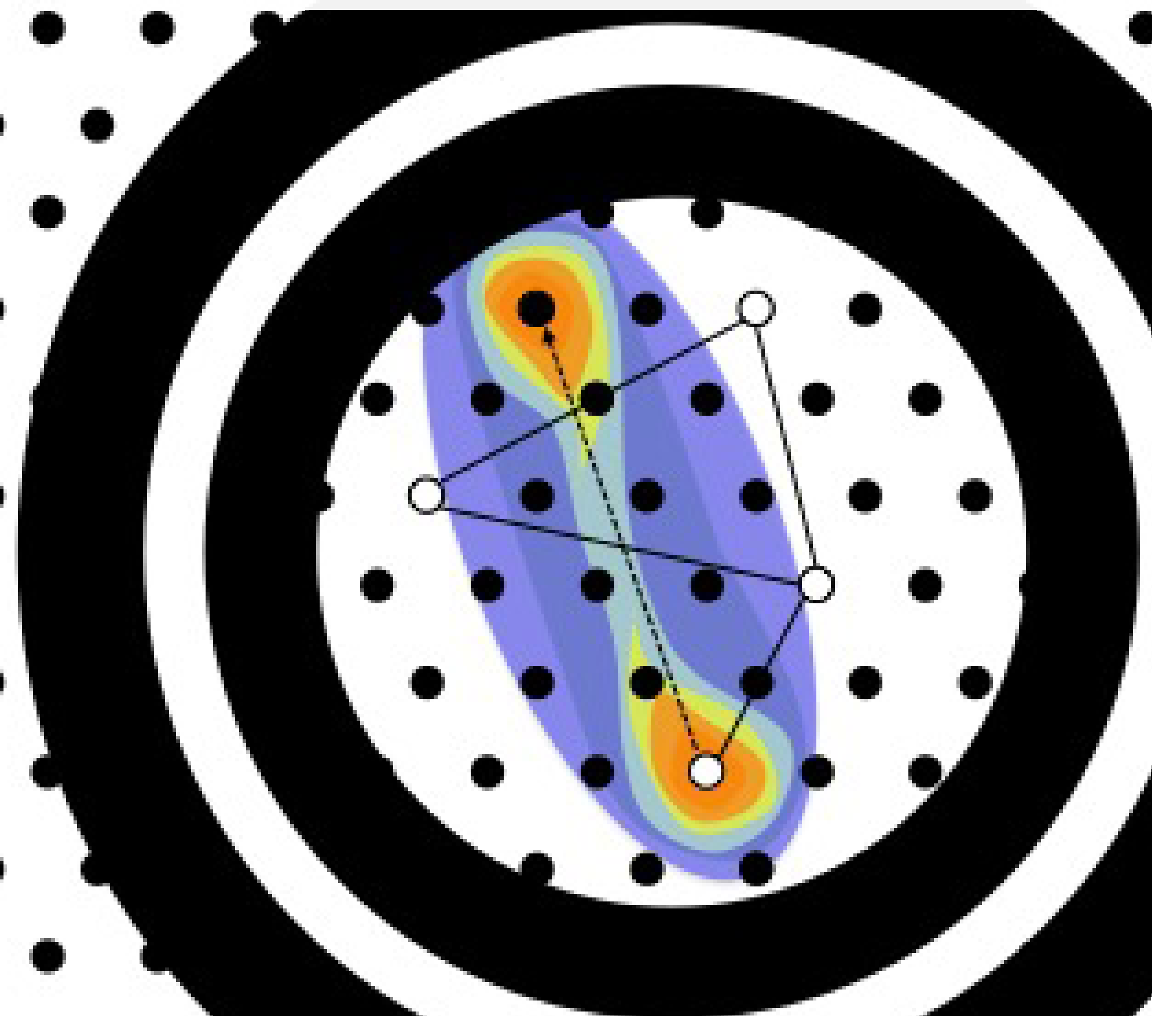
TRAIL THESIS SERIES T2021/16

Lara-Britt Zomer Unravelling Urban Wayfinding

Unravelling Urban Wayfinding

Studies on the development of spatial knowledge, activity patterns, and route dynamics of cyclists

Lara-Britt Zomer



Invitation

You are cordially invited
to attend the public defence
of my PhD dissertation entitled:

Unravelling urban Wayfinding

The defence will take place
on May 6th 2021 at 15h00 in the
Aula Congress Centre at
Delft University of Technology.
Due to current restrictions,
it is only possible to follow
the defence online.

Prior to the defence, at 14:30h,
I will give a brief presentation
about my research.

Lara-Britt Zomer
lbzomer@gmail.com

Paranymphs:
Danique Ton
Ahmed Hussain

Unravelling Urban Wayfinding

Studies on the development of spatial knowledge, activity patterns, and route dynamics of cyclists

Lara-Britt Zomer

Delft University of Technology, 2021

This research is part of the Allegro Project (Unravelling active mode travelling and traffic: with innovative data to a new transportation and traffic theory for pedestrians and bicycles), which is financed by the European Research Council and Amsterdam Institute for Advanced Metropolitan Studies (no. 669792).

Cover illustration by Author.

Unravelling Urban Wayfinding

Studies on the development of spatial knowledge, activity patterns, and route dynamics of cyclists

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus Prof. ir. T.H.J.J. van der Hagen,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op 6 mei om 15:00 uur

door

Lara-Britt ZOMER
Ingenieur Transport, Infrastructuur en Logistiek, Technische Universiteit Delft, Nederland
geboren te Leiden, Nederland

Dit proefschrift is goedgekeurd door de:
promotor: Prof. dr. ir. Serge Paul Hoogendoorn

Samenstelling van de promotiecommissie:

Rector Magnificus	chairperson
Prof. dr. ir. Serge P. Hoogendoorn	promotor
Dr. Oded Cats	promotor
Dr. Ir. Dorine C. Duives	co-promotor

Onafhankelijke leden:

Prof. Dr. Hans van Lint	Delft University of Technology, the Netherlands
Prof. Dr. Dorine Manting	University of Amsterdam, the Netherlands
Prof. Dr. Yusac Susilo	University of Natural Resources and Life Sciences, Austria
Dr. Kari E. Watkins	Georia Tech, United States of America

Reserve lid:

Prof. dr. Marjan Hagenzieker	Delft Univeristy of Technology, the Netherlands

TRAIL Thesis Series no. T2021/16, the Netherlands Research School TRAIL

TRAIL
P.O. Box 5017
2600 GA Delft
The Netherlands
E-mail: info@rsTRAIL.nl
ISBN: 9978-90-5584-291-9
Copyright © 2021 by Lara-Britt Zomer
All rights reserved. No part of the material protected by this copyright notice may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage and retrieval system, without written permission from the author.

Printed in the Netherlands

Acknowledgements

Dear reader,

It has been a great journey and I like to thank you for being part of the experiences, development, and of course the dissertation itself. I am sure I will cherish all special moments for a long time to go, and I hope that many are still awaiting.

Serge, thank you for being such a source of inspiration and the confidence in me. Although I was one of the many PhD's, you always managed to make us feel valued and provided us with the best possible work environment. Oded and Dorine are of course very much part of that environment, and you both have been a big part of my development as a researcher. Thank you for all discussions, support and listening when needed.

Hans, Dorine, Yusac, Kari and Marjan, I am honoured to have you as independent members of my doctoral committee. Thank you for your time. This dissertation would have been different if it was not for Edwin, Peter, Tim and Sascha. Without your help it would not have been possible to collect and analyse the data. Danique and Ahmed, you both had the best possible influence on me, thank you for standing next to my side.

Doing a PhD means being part of an amazing community. I found it overseas, in Delft, Civil Engineering, Transport & Planning, Allegro (and AMS) and room 4.40. Habib and Fiona, you were the best community I can think of in Stockholm. Mignon, thanks for showing me what determination looks like. Danique, Tim, Ilse, Florian, Xiaocheng, Vincent, Marie-Jette, Alexandra, Martijn, Giulia, Alphonse, Yan, Flurin, Yufei, Bernat, Dorine, Victor, Dorine, Oded, Winnie and Serge: you really made me feel part of a great group of people with so many differences and equal similarities. Day-to-day PhD life would have been less fun and interesting without all the coffee breaks, lunches, courses, presentations, conferences. Thanks Niharika, Paul (2x), Malvika, Maria, Xavi, Goof, Pablo, Ding, Yihong, Konstanze,

Alessandro, Marieke, and some more. At the entry of the department you could always find someone to talk to, thank you Moreen, Dehlaila en Priscilla for the candy jar and much more!

Probably many of the best experiences go to Ahmed, thank you for every day since you asked me out. You give me energy and happiness, thank you for always being there. I am grateful for all the people (and cats) that you brought into my life. Thank you mom and Robert for looking after Jopie when we were at conferences (often followed by holidays) and professional advise along the way. Dad, thank you for taking care when my back was hurting and oysters. You all are a big part of who I am today. Jel and Lau, thank you for sometimes (not) asking how the PhD is going. Opa, thank you for being the wise man. To all nieces, nephews, aunts, auncles, and family-in-law, thank you for the great times and laughs during lunches and dinners. You are the best family I can imagine!

Lara Zomer
April 2021

Table of Content

1 Introduction

1.1 Context and Background

Every day residents and visitors find their way through the complex urban network to go to work or get education, or go sightseeing. The density of the urban street fabrique poses more challenges to travellers compared to the more sparse national highway or public transport systems. While it is rare to get lost on your daily commute, it is quite common to deviate from the shortest or fastest route, or to use navigation to avoid congestion and to conveniently find your way. In the Netherlands active modes (pedestrians and cyclists) are accountable for more than 50% of the urban trips. Today, there are still many unanswered questions concerning urban wayfinding behaviour of the active modes. What makes a city easy to navigate, what kind of travel information is easy to comprehend and apply in travel choices, and how can urban design and travel information improve the learned structure of a city?

To answer these questions, we first need to understand the development of the so-called *consideration choice set*. This set consists of all alternatives that a traveller considers prior to deciding where to go and which route to take. It has been argued that the consideration choice set is smaller than the feasible choice set (all possible alternatives, given some heuristic spatial-temporal constraints), and larger than the experienced or observed choice set (only the alternatives at least chosen once during a certain time period). When deciding (how) to move from one place to the next, people base their decisions on their (personal) available spatial knowledge of the city and knowledge of their knowledge of the transport system. Through the interaction between expectation and actual experience, the spatial knowledge evolves with each trip and activity. The underlying theory advocates that with a better understanding of the development of the consideration choice set, urban design and navigation systems can be improved and adapted to better meet the needs and preferences of people.

However, the true consideration choice set is unknown and unobservable from travel patterns nor inferred from experiments or surveys. Adquate methods to achieve this are still lacking, which implies an important research gap. In order to narrow this research gap the aim of this thesis is twofold, analyse how people (citizens and tourists) find the way in urban environments, and identify the role of spatial knowledge in travel patterns. The research objectives require new theories and data, also, models are needed to understand urban wayfinding behaviour and travel patterns.

The context of the research can be explained according to the conceptual framework in Figure 1. Extension of the framework used to investigate the determinants of urban wayfinding styles based on a combination of theories on wayfinding, travel behaviour and the built environment (van Wee 2002; Bovy & Stern 2012). Each trip requires people to make various decisions before and during travelling. These decisions pertain to the modes and routes to be used, and which activities will be performed where and when. Due to individual differences in navigational preferences (e.g. minimize turns and thus choosing a simpler yet longer route) and socio-demographic characteristics (e.g. gender, age, and mode availability) the urban experience differs, and as a consequence, the mental representation of the environment (e.g. perceived accessibility levels, and salient areas) is likely to be different. In turn, these differences will influence the amount of exploration or habitual travelling during future trips. All these characteristics evolve around the *wayfinding attitude* or *style*, defined by *the strategies that people use to decide how to move from one place to another* (Montello 1995). It relates to the set of preferences, selection, and application of navigational strategies, the attitude towards travelling, and the ability to reach the intended destination. As such, differences in travel behaviour are expected to determine the extent to which wayfinding styles and navigational preferences are important to individuals. This dissertation focuses on spatial knowledge development during exploration of a city, hence long-term memory is implicit and thereby reinforcement and memory loss are out of scope (dashed lines).

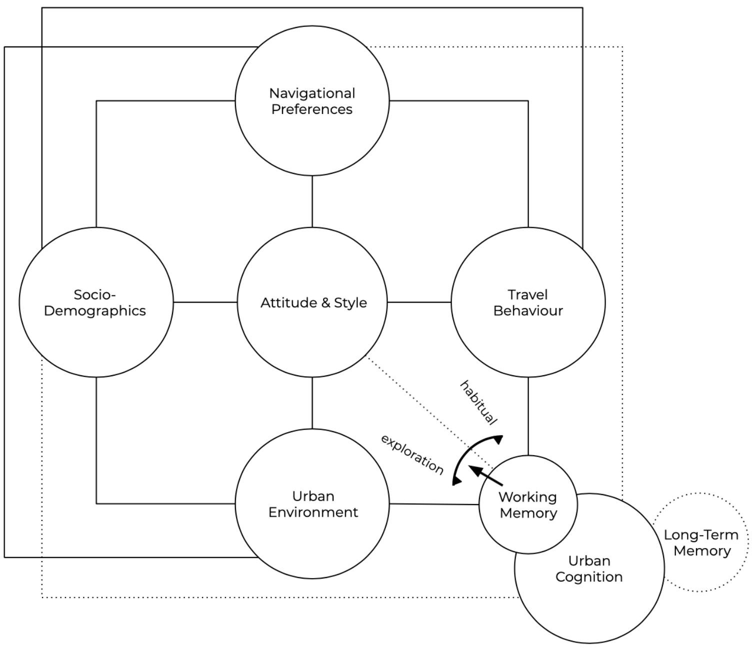


Figure 1. Conceptual Framework.

Understanding how urban wayfinding behaviour relates to travel patterns is important to explain differences in route choice behaviour, to identify difficulties with navigation, for more legible urban planning (Passini 1981; Allen & Golledge 2007), and to provide comprehensible travel information (Dogu & Erkip 2000). However, to date, relations between urban wayfinding styles and the complexity of daily travel behaviour, urban environment, and navigational preferences are partly unknown due to a lack of empirical data.

The remainder of this introduction first details the problem statement (Section 1.1). In Section 1.2, the relation between the research objective and key research questions is explained. The limitations of the research are described in the research scope (Section 1.3). Definitions and concepts provide a theoretic background of related theories in wayfinding behaviour, travel behaviour, and network analyses (Section 1.4). This is followed by the overarching research approach (Section 1.5). The main key scientific and societal contributions of this dissertation are discussed based on the scientific and societal contributions in Section 1.6. Finally, the outline of the thesis is detailed in 1.7.

1.2 Problem Statement

Active modes have been promoted as a sustainable, healthy and inexpensive means of transport that could mitigate urban congestion and urban livability issues due to increased urbanization. Therefore, urban planners and policy makers are looking for ways to create walkable and bikeable cities. Theory, data, and models are needed to understand, predict, and influence activity and movement patterns. However, to provide citizens and tourists with understandable network and travel information, the complexity of human behaviour requires a deeper understanding of how people find the way by foot and bicycle and identification of the role of spatial (network) knowledge. Research efforts are found in the social science research domains, as well as in the more applied, quantitative fields. We argue that combining findings from these areas is crucial to advance the understanding of spatial knowledge acquisition and its impact on travel behaviour.

Over a decade ago, a route choice model has been developed that incorporated conceptually the knowledge acquisition of route attributes based on the mathematical concept of markov chains (Bogers, Bierlaire & Hoogendoorn 2007). A different approach has been introduced by Kazagli, Bierlaire & Flötteröd (2016) aimed to reduce the model complexity by using a network free route choice model based on the mental representation of the environment. In Cenani, Arentze & Timmermans (2013), an activity-based model is presented into how individuals' cognition and mental representation of urban networks develops over time and how the probability of performing a certain activity changes with time. The model takes *perception* (Lang 1987) as an interface between the travellers and their spatial environment, and using *cognition* (Golledge & Stimson 1997) as a way to describe how spatial information is represented in the brain.

In most literature spatial knowledge pertains to landmark, route, and survey knowledge (Freundschuh 1992), detailed in 1.4.2. Based on psychological experiments into children's spatial knowledge development, landmark knowledge is typically the first to be acquired, followed by route and survey knowledge (Siegel & White 1975). However, the dynamics of spatial knowledge acquisition in unfamiliar environments (ie. newcomers or tourists) are hypothesized to be different from children's development and habitual commuters. Instead of choosing from an experienced choice set, they first have to create a choice set, and expend it by exploration behaviour (Golledge 1999). Moreover, landmarks appear crucial in spatial decision-making and can trigger cues indicating turning decisions, or reassuring cues confirming an individual in decisions already made. Although it is known that singularity and saliency are key features of a landmark (Lynch 1960), there are no guidelines for the identification of salient landmarks.

Due to this lack of knowledge existing route choice paradigms are behaviourally inadequate to model mobility choices of tourists and newcomers, as they rely more on the generation of a choice set. This dyad will become problematic in the future as predictions, based on economic property and cheap long-haul travel costs, estimate a growth of 44% to 200%, yielding 28.8 to 41.9 million, tourists in The Netherlands by 2030 (UNWTO 2018; NBeTC 2019). As currently 40% (8 million) of the tourists stay in the capital city Amsterdam this becomes what is called "*overtourism*" when the unequal dispersion of tourists remains. For decades, the main destinations for tourists have been strongly concentrated in a triangle between the Central Station, Vondelpark and Weesperbuurt (e.g.

Jewisch Quarter & Hermitage Museum). Overtourism creates tension between citizens and tourists that decreases the quality of life of both due to excessive noise, nuisance for inhabitants, and pressure on infrastructure (UNTWO 2018). The effect of global strategies and measures to better understand and manage urban tourism heavily depends on the travel behavior of tourists within the respective cities.

The research problems stated above have a fundamental nature, as a comprehensive theory is still absent which can describe when, and why, a particular route or wayfinding landmark is part of a (network) choice set in relation to the (learning of) urban environment, mobility patterns, and information acquisition behaviour of individuals. The main challenge lies in the thorough empirical underpinning and further development and specification of available theories. Not only because of the importance of a strong empirical foundation, but also because of the largely unchartered role of spatial knowledge in travel behaviour modelling. Furthermore, spatial learning modelling of the active mode travellers has received little attention. The corresponding research objective and key research questions are detailed in 1.2.

1.3 Research Objectives

The bicycle is the main mode of transportation in Amsterdam and it is getting more popular amongst tourists. Moreover, the bicycle provides several sustainable, healthy, and inclusive opportunities to disperse tourists to outer areas and alternative destinations within the city (Zomer et al. under review). To advance the understanding of bicycling behavior of tourists, thorough insights are required into activity and movement patterns of tourists and how choices and patterns evolve over time.

The purpose of the study is to narrow the gap between research on travel behaviour research and urban spatial knowledge. Within transport science it is common to estimate and predict travel behaviour using discrete choice or activity-travel models, because of well-defined descriptive and data collection procedures. These methods assume, to a large extent, that decision-making behaviour in travel behaviour is hierarchical and linear. Yet, wayfinding behaviour in cities is defined by the strategies that people employ to decide (how) to move from one place to another within an urban area (Montello 1995). Regarding the understanding of urban wayfinding, a theory is still absent which can describe when, and why, a particular route or wayfinding landmark is part of a (network) choice set in relation to the (learning of) urban environment, mobility patterns, and information acquisition behaviour of (active) travellers. Closing this gap is necessary to develop an experimental platform to test innovative information services in different (urban) scenarios prior to deployment.

To this end, it will be studied how travel behaviour, urban environments, and information services impact spatial knowledge development. These insights into the dynamics of the internal representation can be used as additional inputs for adapted activity-travel models and microscopic simulations. In order to develop theory, conceptual and mathematical models on the development of active modes' spatial knowledge in activity-travel modeling across urban environments, the research objective is:

Unravel the role of spatial (network) knowledge and particulary how active mode travellers find their way in urban environments.

The main objective can be divided into four research questions:

1. *What are the relevant dimensions for characterizing urban wayfinding styles, and how do they relate to daily travel behaviour? (Chapter 1)*
   The first paper draws on theory testing based on existing literature on wayfinding behaviour and investigate the relation with travel diary data.

2. *How can open spatial data be used to identify salient and legible urban areas (landmarks)? (Chapter 2)*
   The second paper presents a methodology based on spatial analytics to use open spatial data to characterize salient and legible areas in an urban environment that are presuminlgy more easy to memorize.
3. *What is the relation between the spatial and temporal activity patterns of visitors applied on tourists by bicycle in the metropolitan area of Amsterdam? (Chapter 3)*
   The third paper provides new insights into activity patterns of tourists based on a large empirical field study of GPS trajectories of bicycles. The insights are used to develop new theories to better understand and influence travel behavior of tourists by bicycle in crowded cities.
4. *What determines the spatial boundaries of the route selection space of tourists travelling by bicycle, and how does spatial (network) knowledge acquisition influence the movement pattern to the next activity? (Chapter 4)*
   The fourth paper draws upon the same GPS trajectory data of bicycles of tourists as the third paper. However, now the spatial characteristics of the route patterns of tourists are used to describe the (development of) spatial knowledge of tourists. A model is used to estimate to what extent the detour ratio and deviation area of a bicycle trip can be predicted based on the theoretically acquired spatial knowledge.

The research approaches for each sub-objective and how they relate to individual studies are further detailed in Section 1.6.

1.4 Research Scope

The overarching goal of this dissertation research has been described in the Allegro research proposal and finds its origin in the assumption that individual and collective behaviour of active modes differs strongly, perhaps even fundamentally, from motorized vehicles and public transport. It has been hypothesized that the lower travel speed of active modes impacts the perception of salient waypoints, which in turn influence spatial knowledge development (of the consideration choice set) and ultimately the flexibility in choice options, e.g. route choices. A deeper understanding of human behaviour is necessary to explain the complex behavioural dimensions and interactions of pedestrians and cyclists in an urban context. Therefore, the overarching goal of Allegro is:

> *"To develop and empirically underpin comprehensive behavioural theories, conceptual and mathematical models to explain and predict the dynamics of pedestrians, cyclists, as well as mixed flows at all relevant behavioural levels, including acquiring spatial knowledge, activity scheduling, route choice and operations, within an urban context, with a special focus on the role of ICT on learning, and choice behaviour."*

Within the Allegro thesis series, this dissertation research has a focus on the behavioural level of *spatial knowledge acquisition* of cyclists. The research scope is to advance theories, conceptual models and mathematical models on the development of active modes' spatial knowledge based on the urban movement patterns. The exploratorive work should provide an empirically underpinned foundation that will support with understanding how people acquire and represent knowledge about the environment they are travelling through. In doing so, we unravel the relation between urban wayfinding strategies and travel patterns, identify urban salient areas, explore the urban activity pattern of travellers with limited spatial knowledge, and we model spatial route dynamics as a function of the development of experience and spatial knowledge. Combined, the findings of this dissertation aim to understand more about the development of spatial knowledge, and particulary the relation with wayfinding of active mode travellers in urban environments.

In a broader context, the knowledge, sometimes referred to as the *consideration route* and *activity choice set*, determines the *activity-travel level*, i.e. how travellers schedule activities, choose where and when to perform what activity, as well as, choose the routes towards these locations. The relation between knowledge acquisition and activity-travel level provided opportunities for collaboration in data collection as well as research efforts, together with Danique Ton (entitled "Unravelling mode and route choices of active users") and Florian Schneider (entitled "Unravelling trip chaining behaviour of active users"). This also implies that the relation between operational behavioural (e.g. gazing and perception), spatial knowledge acquisition, and mode choice behaviour is outside the research scope of this dissertation.

The research on activity patterns and route dynamics within the dissertation focuses on cyclists. To generalize findings to active users in general, similar research approaches can be applied on pedestrian data. Based on the assumption that the travel speed differences between active users and motorized vehicles are likely to result in fundamentally different urban wayfinding behaviour, in the respective studies the cycling patterns are compared with motorized vehicles.

The majority of the existing studies that investigated the relations between learning behaviour, wayfinding, and travel patterns, are in controlled (fictive) or small scale environments. These findings do provide some insights in the complexity of the behaviour, and even investigate the role of ICT. However, if similar processes govern the daily commute patterns or exploration patterns of tourists or newcomers is unknown. Therefore, the aim of this research is to conduct an explorative research to unravel the spatial learning process through wayfinding and the relation with travel behaviour, and investigate the link between knowledge acquisition and tactical decision making of travellers.

1.5 Definitions and Concepts

This section provides an overview of the state-the-art definitions and concepts related to wayfinding strategies, spatial knowledge, travel characteristics of cyclists, and dynamics and development used in the state-of-the-art. Chapter 1 combines wayfinding strategies, spatial knowledge, and daily travel behaviour. Chapter 2 operationalizes spatial landmark knowledge. In chapter 3 travel behaviour theories on activity patterns are used to formulate hypotheses. Chapter 4 investigates the route dynamics and spatial knowledge development.

1.5.1 Wayfinding strategies

Urban wayfinding behavior is defined by the strategies that people use to decide how to move from one place to another within a city (Montello 1995). It relates to the preferences, selection and application of navigation strategies, the attitude towards travelling, and ability to reach the intended destination. While travelling through the urban environment spatial knowledge will be utilized, acquired, and memorized. Based on small-scale environments three types of wayfinding systems have been identified: egocentric, allocentric, and map-like orientation and navigation (Piaget 1968; Stea and Blaut 1973).

1. **Egocentric** (based on self-centred) **orientation**. Spatial orientation using axes or planes with respect to one's own body in order to orient where one is within the environment. For example landmarks or street names signs within the visual field provide information about the local whereabouts.
2. **Allocentric orientation**. Provided a self-centred orientation, there exists also orientation towards destinations not within the direct perceptual field. Within local areas, these destinations can be related, but not as a sense of the whole. For example, while travelling from work to home one is aware of the direction to the origin and to the destination. Additionally, the direction to the shopping area close to home is known from home, but not from work.

3. **Geocentric** (based on coordinated) **orientation**. Regardless of self-centred orientation, using cardinal directions the urban environment is understood by the relative directions between locations. As such, wayfinding between work and the shopping area close to home can be done without much trouble. Note that it is not necessarily based on real distances, rather upon physical features in the experienced environment. Geo-centric orientation is believed to be important to understand maps and communicate directions.

However, it remains questionable to what extend this classification is also meaningful in large-scale environments. Based on research about wayfinding behaviour of animals (i.e. mammals, ants and bees) a related concept exists to find the way using the coordinates of origins and destinations (Richter and Winter 2004). Path integration estimates the current position and provides direction and distance to the origin, regarding the original three types, Path integration can be seen as a transition between allocentric and geocentric wayfinding behaviour. Still more realistic learning and memorization processes can be incorporated to their approach, as without prior knowledge the spatial memory is a tabula rasa and effort is minimized by a goal-seeking strategy to first explore the most proximate location(s).

1.5.2 Spatial knowledge

A second element that is important to find the way in large-scale spaces is environmental cognition. That is, space must be cognitively organized and memorized when the entire route cannot be perceived at once, or when all feasible routes cannot be perceived as a sequence of discrete views (Stea & Blaut 1973). Environmental cognition consists of spatial knowledge of locations (distance, direction, and relative relation) and associated (descriptive and evaluative) attributes. The latter, namely associated attributes, are dependent on the measurement scale (e.g. country, city, shopping mall). There is a long-lasting hypothesis without consistent evidence that assume knowledge of individuals' cognitive maps can be used to predict the spatial behaviour (Fishbein 1967).

Urban spatial cognition is considered to be the internal (personalized) knowledge representation of the urban environment in our mind. The internal representation consists of both spatial and temporal dimensions, as no feature can be experienced as if it were a stand-alone item, each feature will always be experienced in relation to its contextual surroundings (spatial position). Cognitive sciences distinguish three levels of knowledge that can be acquired (simultaneously) in time (Siegel and White 1975).

1. **Landmark knowledge**. Information about location of objects in space. A landmark can be used as a crow flight direction for navigation, orientation on changes in direction or to maintain course (McNamara et al. 2008). Aggregate urban landmarks can been seen as salient urban areas that possess no-ticeable characteristics that make them distinct from their surroundings. From a theoretical perspective, a landmark is salient (distinct) in relation to its immediate surrounding or context at large. Salient urban areas are considered unique, either because of dissimilarities to their (local) area, and/or else, because of characteristics considered similar in comparison to other (global) areas. Presumably, the more distinctive a landmark or area, the easier it will be to memorize and incorporate this saliency into the spatial route knowledge to be drawn upon in future. Therefore, salient urban areas are hypothesized to be important to structure spatial knowledge in long-term memory (Couclelis et al. 1987; Sadalla et al. 1980; Montello 1997). It appears, whereas in urban planning, landmarks appear firmly grounded concepts, their appliance to large-scale environments is cumbersome, particularly, when buildings are unequally distributed. Based on Lynch (1960), regarding their identification, generally, landmarks are analyzed as geo-referenced points or buildings;
2. **Route knowledge**. Sequences of landmarks associated with a process of decisions and actions;
3. **Survey knowledge**. A broader understanding of the urban environment that can be used to construct routes to unseen landmarks or locations or to construct alternative routes.

Clearly these three "levels" are not distinct, which makes operationalization complicated. Approaching cognition from neurosciences, Manning et al (2014) based on their model more than ten distinct categories using linguistic classification of reports from taxi drivers while watching a video of their navigation performance through a virtual London (Spiers and Maguire 2006). However their approach focuses on procedural knowledge (action planning, expectations, spontaneous route planning etc.).

1.5.3 Travel behaviour & patterns

Travel characteristics are often described in distances, duration, and frequency. Traveling is often seen as a means to get to another destination where another activity can be performed compared to your current location. However, traveling in areas with a high quality (nature or architecturally) and the healthy benefits of active modes like walking and bicycling, also may give traveling a positive utility (Anable and Gatersleben 2005; Ory and Mokhtarian 2005; Steg 2005).

Understanding travel behavior is dependent on the activities in which individuals like to participate at their destination(s). As well as, while traveling, and the options they have to fulfill partaking in the activity and arriving at the desired destination.

The questions studied in travel behavior are broad, and are probed through activity and time-use research studies, and surveys of travelers designed to reveal attitudes, behaviors and the gaps between them in relation to the sociological and environmental impacts of travel. To determine which factors influence individual travel behaviour often descriptive methods are used, such as a travel diary, often part of a travel survey or travel behavior inventory. Large metropolitan areas typically only do such surveys once every decade, though some cities are conducting panel surveys, which track the same people year after year.

1.5.3.1 Daily travel behaviour

The *mobility portfolio* describes the amount of travelling per mode, which can described by travel distance, time and number of trips. The daily mobility patterns pertain to both mode choices and preference hierachy towards different modes. Travel diaries provides three possible indicators per mode; distance, travel time, and number of trips (De Haas et al. 2018). As there are significant differences in travel distance and time per mode, We average the number of trips per day as reported in the three-day travel diary, to identify the daily mobility pattern (Ton et al. 2019).

1.5.3.2 Activity patterns

Network analyses aims to reveal the topological features to understand the dynamics of the activity and route network resulting from the activity and movement patterns. Of interest is to investigate if there are so-called communities, i.e. set(s) of activity zones (routes) that are generally visited in combination on one day. The existence and behavior of communities will influence where wayfinding systems, and which content, should be located to a) stimulate tourists to remain in a specific community, and b) distribute tourists to other, less crowded, communities. Differences can be observed in the behavior of communities in terms of degree, clustering coefficient, betweenness centrality (Newman 2006).

1.5.3.3 Route patterns

The spatial behavior of tourists is a direct function of their experience considering the built environment (Golledge & Stimson 1987). To the best of the authors' knowledge, the development of spatial behavior when travelling in an unfamiliar environment is largely unknown, especially when these movements are performed by bike.

Spatial choice sets are the result of a complex interplay between spatial restrictions, activity space, and personal abilities and preferences (Bovy & Stern 1990; Manaugh & El-Geneidy 2012). The origin of the spatial route choice set concept can be found in Hagerstrand's space-time geography

(Hägerstrand 1953). The potential path area (PPA) is the projected ellipse of the space-time diagram on the surface, which represents all locations that a person can occupy during the available time between two sequential activities (t_i, t_{i+1}) (Miller 2005). What the potential path area (PPA) represents at trip level, is the spatial route choice set at individual level. More common in literature, and closely linked to the spatial route choice set, is the activity space.

Similar approaches have been used to represent individual and household activity spaces, for instance using ellipses (Newsome, Walcott & Smith 1998), minimum spanning trees and kernel densities (Schönfelder & Axhausen 2002) and local travel index (Manaugh & El-Geneidy 2012). The model results and significant determinants of these four studies are documented in Table 1. Schönfelder & Axhausen (2002) reflect on these methods, and conclude that activity space ellipse overgeneralizes the spatial pattern leading to an oversized area, kernel densities ignore connections between activity locations, and minimum spanning tree only captures the spatial distribution of the activities. They propose to combine the minimum spanning tree with a spatial buffer to incorporate the size of human activity spaces, called the road network buffer approach.

Only one study has analyzed the existence of the route selection space (RSS) based on a large data set of car drivers, which was coined the boundary of human routes (Lima et al. 2016). They found that 95% of all detours are bounded by an ellipse. This ellipse can be described using the eccentricity, which is the deviation between geodesic trip distance and the maximum value of the sum of the two geodesic distances between the origin and destination, and each point along the trajectory. They compared their findings with eccentricities from optimal routes and concluded that human routes have wider spatial route choice sets. Furthermore, they found indications that the RSS of car trips is independent of the Euclidean trip distance.

Next to the network layout, also other variables have been identified to impact the RSS. For instance, Bovy & Stern (1990) hypothesize that subjective spatial restrictions, personal preferences, and activity patterns determine the boundary of the RSS, leading to individual route selection spaces, while Golledge & Stimson (1987) developed a theory that demonstrates that spatial behavior of people is a direct function of their individual experience with the built environment. Yet, the dynamics of the individual route selection space when familiarity is under development and the relation with travel behaviour are currently undetermined.

1.5.4 Route dynamics and spatial knowledge development

A meta-literature review conducted almost three decades ago identified differences in relative accuracy of cognitive distance. Immediate distance observations are on average 8% higher, while memorized previously visited destinations and inferred distances to unknown destinations are 9% and 25% smaller compared to the actual direct distance (Wiest & Bell 1985). Another research direction analyses how the acquisition of internal spatial representation of cities while training to become a licenced taxi driver relates with gray matter volume in the posterior hippocampus changes to memory profiles (Woollett & Maguire 2011).

Familiarity can evolve with every trip and activity and affect activity patterns and route choices. However, spatial and network knowledge is only acquired when experiences of previous trips and activities are processed and memorized. Moreover, the perception of attributes improves when the acquired knowledge is appropriately applied to future and new activity and route choices (Stern & Leiser 1988). The ability to process and apply the newly acquired knowledge (directly) to future trips depends also on individual spatial abilities and preferences. Diminution and memory loss or selection ensures that excess information is lost and important features are retained (Miller 1956). Limited memory retention has been modelled in a cognitive learning model of daily activity-travel patterns based on the shortest path and attention and sensitivity to environmental attributes (Cenani, Arentze & Timmermans 2012).

Based on the above literature it is assumed that spatial behaviour depends on direct distance between origin and destination and trip purpose characteristics. Unfamiliar travellers start without spatial or network knowledge of the urban environment, but according to the accretion principle their familiarity develops already after the first trip (Stea & Blaut 1973). While the experience of unfamiliar travellers can be quantified by means of number of trips, historic travel experience, and previously acquired spatial knowledge and routing behaviour. To find the way to the next activity location the acquired knowledge can influence the characteristics of the next trip. This depends on trip length, size of new and old areas that have to be explored and retraced, and time pressure.

1.6 Research Approach

According to the postpositivist philosophy empirical studies on human behaviour do not allow to deduce the absolute truth of knowledge, as the values and knowledge of the researcher affects study results and outcomes (Phillips & Burbules 2000). Therefore, a transformative mixed method procedure has been used to develop and collect quantitative and qualitative empirical datasets to explore new theories to model spatial knowledge acquisition based on spatial analytics, urban cognition, and travel behaviour. The four studies that form the backbone of this dissertation are all based on state-of-the-art backgrounds containing a review of the most relevant literature on urban wayfinding attitudes and styles, salient areas, and activity patterns and routing behaviour of tourists travelling by bicycle. The associated research approach is visualized in Figure 2.

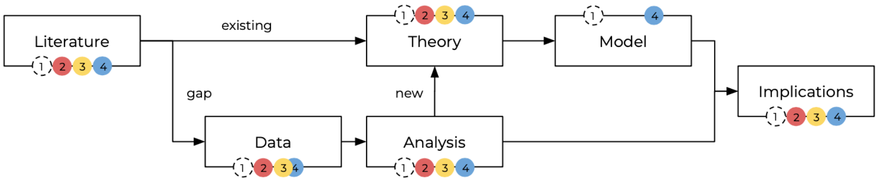


Figure 2. Research Approach.

1.6.1 *Data collection methods*

At the time of the research proposal was written, the individual and collective behaviour of active modes appeared data poor. In particular, we lack high quality data of cycling behaviour. Consequently, as empirical research the main goals were to collect, process, and analyse data that would provide innovative insights in the complex behavioural dimensions of active modes. Table 1 shows a advantages and limitations of possible data collection methods to unravel wayfinding and/or travel behaviour. Stated preference and virtual reality experiments both allow for hypotheses testing in controlled environments (Skov-Petersen et al. 2018; Vilar and Rebelo 2008). Cognitive data pertains to EEG, FMRi, and eye tracking and provide insights into knowledge p rocessing (Hartley et al. 2003; Kiefer, Giannopoulos and Raubal 2014). Questionnaires are the-go-to approach that is often combined with other data collection types. It can range from psychophysical experiments to attitudinal surveys (Zacharias 1997; Hegarty et al. 2002). Travel diaries have been a core pillar in travel behaviour research as they provide comphrensive, yet detailed, information on activity and trip level (Kirasic 2008; Schönfelder and Axhausen 2016). Various revealed data collection approaches are gaining popularity due to low respondent burden, limited scaling issues, and possibility to provide fine grained and high quality data to measure travel patterns. Popular tools are GPS trackers, mobile phones and other electronic devices, and cameras.

Based on the advantages and limitations of data resources to unravel wayfinding and travel behaviour two approaches seemed viable. A combination between a stated preference study in a virtual reality environment gives the potential to simultaneously collect cognitive data (EEG and eyetracking). Follow-up questionnaires could be used to gain insights into the perception of respondents. This option could provide innovative insights into the role of ICT. However, as there was not a strong emprical foundation of the theories in literature, the design of such an experiment would be compromised as many assumptions would be implicit. Therefore, another direction has been chosen; a combination between travel diaries of a representative sample of the Dutch population with an additional survey targeted at preferences, attitudes, and wayfinding styles of active modes and revealed GPS data of tourists that rented a bicycle during their stay in Amsterdam. By conducting this research and associated data collection efforts, empirical insights will be gained that could provide input for future studies where the role of ICT and social interactions can be investigated in more controlled environments. More details on the recommendations for future research can be found in 6.2.

Table 1. Advantages and limitations of data concerning wayfinding and travel behaviour.

Data	Stated Preference	Virtual Reality	Cognitive Data	Questionnaires	Travel Diary	Revealed Data
Avantages	Allows to investigate new situations. Controlled environment. Realism is questionable. Well defined analysis methodologies.	Allows to investigate new situations. Controlled environment. Realism is questionable. Possibility to study a larger choice set size as less emersion abilities are required compared to SP.	Unique insights in relation between brain activity and perceptional abilities. Useful to investigate the link between operational and strategic level.	Commonly used within both travel behaviour and wayfinding research. Flexibility. Low effort to combine with other data types. Insights on attitudes, perceptions.	MPN provides a unique representative sample of the Dutch population, including additional questionnaires. Structured way of data processing, and easy to analyse using statistical software.	Commonly used within travel behaviour research.
Limitations	Measures controlled behaviour. Depending on the design it can be time intensive for respondents. Requires emersion abilities of respondents, which is limited. Respondent likely to exeggerate economical decision-making due to systematic presentation of information.	Measures controlled behaviour. Design of a good controlled environment is time intensive. Depending on the design it can be time intensive for respondents.	Fluctuations in data require sophisticated cleaning procedures. Sensitivity of equipment requires dedicated rooms to conduct experiments. Time intensive for researcher.	Reporting behaviour subject to human errors. Quality of the data depends on quality of the questions and design of the questionnaire.	Reporting behaviour subject to human errors. Time intensive for respondents.	Erronous and sparsity in data require tailored cleaning procedures. Privacy regulations.

More specifically, travel diaries (MPN) are combined with a new questionnaire on perception and wayfinding behaviour of active modes (PAW-AM) to identify urban wayfinding styles and analyse the relation with travel behaviour, navigational preferences, urban environment, and socio-demographics. Second, Open geospatial data from the Municipality of Amsterdam is processed to identify and analyse the spatial distribution of legible and salient areas locally (beacons or landmarks) and globally (neighbourhoods). Finally, a large-scale data collection study is designed to unravel the spatial **L**earning process of, and **U**nderstand the impact on **CY**clists' behaviour (LUCY). Two studies are based on fine-grained GPS data of activity and routing patterns by bicycle of unfamiliar people (tourists) within an urban environment.

1.6.2 Data analysis methods

The causal relationships from existing literature and that have been hypothesized from the processed data are analysed to refine and develop theories that describe how spatial learning affects travel behaviour in cities.

1.6.2.1 Mobility portfolio and wayfinding styles

In line with Kaufman and Rousseeuw (1990) and Everitt (1993) *latent class* (LC) *cluster analysis* is a method to classify people in mobility patterns based on their reported travel behaviour, when both the number of different mobility patterns and its properties still need to be determined. An important difference between standard cluster analysis techniques and LC clustering is that the latter is a model-based clustering approach (Vermunt and Magidson 2002). This means that LC analyses are statistical probabilistic clustering methods, each traveller is assumed to belong to only one mobility pattern, while the uncertainty of the class membership is taken into account. Posterior class-membership probabilities are computed from the estimate model parameters and observed scores. There are three benefits of LC cluster analyses to unravel mobility patterns are i.) the flexibility to use both simple and complicated distributional forms for the observed variables (e.g. number of trips), ii.) restrictions can be imposed on parameters to obtain more parsimony, which can also be validated, and iii.) scaling of the observed variables is irrelevant when using normal distributions.

Exploratory factor analysis is a statistical method to explore the underlying relation between a large number of measured variables, which are assumed to be related to a smaller number of "unobserved" factors (Fabrigar et al. 1999). In order to reduce the dimensionality of 23 questions of the self-report questionnaire (Santa Barbara Sense of Direction) an Exploratory Factor Analysis is used to derive a set of lantent constucts that represent the urban wayfinding styles. Principal component extraction and varimax rotation have been applied to minimize multicollinearity effects and to identify the underlying dimensions of urban wayfinding styles.

1.6.2.2 Salient areas

An iterative *grouping analysis* in ArcGIS is used to explore the reliability of the determinants to identify different urban morphologies. The goal is select the metrics (mean, standard deviation, minimum and/or maximum are most meaningful and reliable to decsribe an 100x100 metres area in terms of building volume, building surface, number of floors, number of buildings.

Cluster and outlier analysis is applied in many domains, such as economics and geography to identify concentrations of values and outliers that explain (behavioral) patterns (Anselin 1995). To identify salient areas, in ArcGIS the cluster and ourlier analysis is based on Anselin Local Moran's I, using the selected determinants as input fields. This analysis is often preferred over hotspot analysis based on the Getis-Ord Gi*, as it also identifies statistically significant spatial outliers, which are expected to be the most important aggregate urban landmarks. An inverse distance squared is used because nearby neighboring grid-cells have a much larger influence than grid-cells further away.

The Gini coefficient according to Brown's formula has been used as a comparative measure of dispersion relative to salient urban areas within Amsterdam. This analysis is preferred over the multi-distance spatial cluster analysis because it is scale dependent (Tsai 2005). The ratio analyses are used to measure the inequality of the distribution of salient urban landmarks in Amsterdam, based on 1.) the extent to which an urban area is salient, and 2.) the number of salient urban areas within a certain distance field of a salient urban area. For example, a distance field of 300 metres represents 8 grid-cells surrounding a salient urban area. The Gini coefficient can range between 0 and 1, with 0 representing perfect equality, and 1 representing perfect inequality of the distribution of salient urban areas in Amsterdam.

1.6.2.3 Activity patterns

K-means clustering has been used to derive activity zones from GPS data points classified as activity locations to locate the main tourist destinations by bicycle in Greater Amsterdam Region. The benefit of this unsupervised algorithm is that the resulting activity zones are solely based on the spatial proximity without a reference outcome. Generally, the Euclidean distance is used in spatial k-means clustering analysis. In this case, however, an Euclidean distance measure would provide unrealistic clustering results, given that most tourists by bicycle diverge from the direct (Euclidean) line between the identified activity location (of the parked bicycle) and the main destination (at the activity zone) due to street patterns in Amsterdam. A Manhattan (city block) distance computes the absolute differences between coordinates of pair of objects (Kaufman & Rousseeuw 2009), thus providing a more realistic clustering result. To ensure the avoidance of local minima 90 initializations are used. The number of clusters is determined based on the minimum number of clusters where there exists a peak at mean silhouette value (i.e. the consistency of points within each cluster) compared to neighboring clusters (Rousseeuw 1987) and the value of improvement of Best Total Sum of Distances, which should be higher than the average where the line stabilizes. However, activity locations are unevenly distributed among Amsterdam city center and outer areas. Thus, two clustering procedures are performed, one for the locations inside the ring road and one for the locations outside the ring road.

With a *network analysis* we aim to reveal the topological features to understand the dynamics of the activity zones network resulting from the activity pattern. Of interest is to investigate if there are so-called communities, i.e. set(s) of activity zones that are generally visited in combination on a given day by tourists of TSH. For instance, if the majority of the tourists that visit(ed) Museum Square also visit(ed) the Vondelpark followed, or preceded, by Leidse Square on the same day the three activity zones are likely to belong to the same community. The existence and composition of communities will influence where wayfinding systems, and which content, should be located to i) stimulate tourists to remain in a specific community, and ii) distribute tourists to other, less crowded, communities. Differences can be observed in the characteristics of communities in terms of weighted degree, clustering coefficient, and betweenness centrality (Newman 2006). First, the number of communities in the tourists' activity zone network can be derived, based on the maximal modularity.

The aim of the *activity space analysis* is to identify spatial differences between activity communities. Activity space depicts the area where activities are performed by an individual (Newsome, Walcott & Smith 1998). The activity space of commute behaviour is often based on activity chains (primary activity, i.e. home - secondary activity, i.e. grocery shopping/pick up - primary activity, i.e. work/education, ... - primary activity, i.e. home) and used to identify the area where activities are likely to be performed considering time and spatial constraints. Tourists activity behaviour is presumingly less hierarchical compared to commuters because mandatory, preplanned activities, such as work or education, are rare. In this dissertation activity locations of each tourist day have been used to determine the revealed activity space . Therefore, the convex hull is used to compute the Euclidean space surrounding the activity locations a tourist chooses to visit on a given day. The aim is to analyse if the revealed activity space and corresponding activity pattern are significantly different depending on visiting activity zones within the city (Central Station) or in the outer areas of Amsterdam (Zaandam Region).

1.6.2.4 Spatial knowledge acquisition

The spatial route choice set of an individual tourist is latent. Yet, provided with many trips, the spatial probability distribution can be estimated. The spatial probability distribution can be used to analyze the dependency of the route selection space on amonst others spatial knowledge acquisition, urban street network, and travel mode. The spatial probability distribution consists of a bivariate histogram plot of X and Y that visualizes the route selection space of all trips in a normalized space with respect to scale and direction. This requires a transformation of Euclidean trajectory coordinates for each *trip*$_i$ of *tourist*$_n$ to Cartesian coordinates, with all origins at location (0,0) and destinations at location (0, 1). The bins contained in the bivariate histogram plot represent the relative number of observations, as such the sum of all bins equals 1.

1.6.3 Modelling

Generalized Linear Models and Generalized Estimating Equations are used in two studies to assess the relation between determinants of urban wayfinding styles and spatial learning affecting the route selection space, while controlling for correlations. Moreover, the aim of this dissertation is to discuss the implications of the new theories and findings to active-modes policy, urban design, and travel information.

The goal of this dissertation is to answer the research question "To what extent can differences in the urban wayfinding styles coined as 'Orientation Ability' and 'Knowledge Gathering & Processing Ability' be explained by a comprehensive model including the relations with socio-demographics and motility, urban environment, navigational preferences, and travel behaviour?" This can be investigated using different statistical models, including *Generalized Linear Models* (GLM) (Nelder & Wedderburn 1972; Diggle, Liang & Zeger 1994; Cox, West & Aiken 2013) and multinomial logistic regression. One of the major pitfalls of multinomial logistic regression is the reduction in degrees of freedom when many parameters are included. Different from regression models, GLM assumes that there is no clustering of the data and thus responses of all respondents are mutually independent.

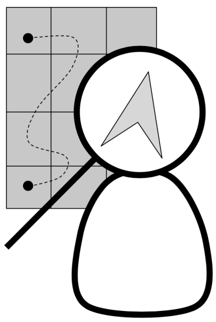
Generalized Estimating Equation (GEE) models are used to assess if the movement patterns of tourists become more efficient when the familiarity with the built environment grows, which leads to a decline in detour ratio and maximum deviation and decrease of eccentricity and increased efficiency of the curvature. GEEs are an extension of GLMs, and are developed to analyse longitudinal and/or correlated data (Liang and Zeger 1986). This approach is conceptually different from multilevel and hierarchical models as GEEs do not explicitly model the variation. Instead it focuses on, and estimates the similarity of the observations (Hanley et al. 2003; Ballinger 2005). As a result GEEs are marginal models, they model a population average. The results should be interpreted as with every unit increase of an explanatory variable across the population identifies the change in the average response of the dependent variables corresponds.

1.7 Scientific and Societal Contributions

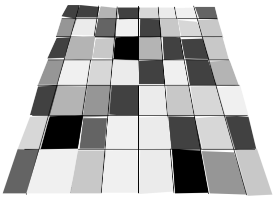
The general contribution of this dissertation is the increased understanding of the role of spatial knowledge and particularly how active mode travellers find their way in urban environments contributes to science and society.

1.7.1 Scientific Contributions

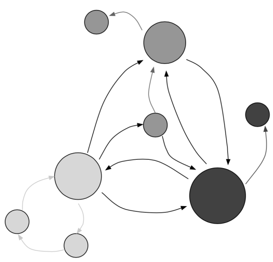
The contributions to science are based on empirical and experimental data, the developed cognitive models will focus on the development of active modes' spatial knowledge regarding urban environments over time. The specific scientific contributions of this thesis can be grouped in four different categories, namely urban wayfinding styles, salient areas, activity patterns, and route selection spaces.



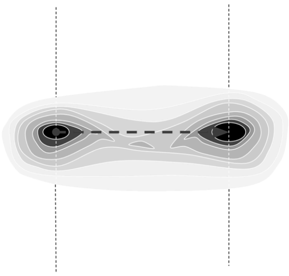
- Provide theoretical insights of how urban wayfinding behaviour relates to daily travel patterns.
- Empirical insights into the combined effects on two identified wayfinding styles have been investigated with a large and representative empirical analysis using an exploratory factor analysis and Generalized Linear Models (GLM).
- An objective and critical evaluation of the GLM results based on contingency tables and confusion matrices.
- Discussing the possibilities and relevance of wayfinding styles for route choice behaviour.
- Results provide evidence that predominantly different processes describe each wayfinding style.



- Spatial analytic methodology to handle open-source datasets to identify urban wayfinding landmarks as salient urban areas.
- Salient urban areas are identified by building volume, surface, height, building year, and the number of buildings in a 100 square metres grid-cell.
- Findings have been applied to identify differences in distribution of clustering and dispersion between local and global salient urban areas using the Gini coefficient in Amsterdam Metropolitan Region.



- Empirical data to unravel spatial and temporal characteristics of tourist activity patterns by bicycle in Amsterdam Metropolitan Region.
- Activity detection algorithm has been developed to process GPS into 10,347 activity locations and 105 zones of 1,817 unique tourist day pattern.
- Spatial relations between activity zones are analyzed based on a network analysis that indicates the influence of the location of hotels on activity patterns.
- The relation between activity space, compactness and travel time ratio provide insights into the spatial distribution of tourists.



- Provide insights into travel choices of city tourists travelling by bicycle using fine-grained GPS data of 1,810 tourists making 8,490 trips in and around Amsterdam.
- Operationalisation of the route selection space dynamics, and perform analyses based on spatial probability distributions and Generalized Estimating Equations (GEE).
- The findings show that route selection space of tourists depends mainly on trip purpose.
- The findings also show that tourists learn within a day though the number of trips and new activities.

1.7.2 Societal Contributions

The societal contributions of this thesis can be grouped in three different categories, namely policy, urban design, and travel information.

1.7.2.1 Policy

Discussing the possibilities and relevance of wayfinding styles for route choice behaviour, provision of comprehensible travel information.

Based on findings in literature, it can be expected that respondents with better sense of orientation choose routes with shorter travel distance and time, but not necessarily higher travel speed. This requires flexible navigational preferences as the structure and layout of each urban environment demands different abilities. However, both GLMs did not include navigational preferences to minimize travel distance or time. Regarding the provision of comprehensible travel information, this indicates that wayfinding styles are more related to number of turns, bearing line and short-cuts than travel distance or travel time. In the future, a similar study including travel data at route level could be used to investigate differences in route choice behaviour and variability.

1.7.2.2 Urban design

Discussing the possibilities and relevance of wayfinding styles for design of legible cities.

Based on the models it seems that a combination of high Orientation Ability and Knowledge Gathering & Processing Ability will correspond to higher variability in the streets of chosen routes. With higher (perceived) connectivity of the bicycle infrastructure more Orientation Ability is required than average. This implies that people with lower levels of Orientation Ability will compensate for the complexity of the urban wayfinding task by preferring a longer route along familiar streets. Thus, even if high connectivity exists, but all people have low orientation abilities, still not much route variation will occur and it will become more difficult to mitigate congestion and distribute large cyclists flows more evenly. Insights related to navigational preferences and urban environment on Knowledge Gathering & Processing Ability can be interpreted as for people that do not wish to make short-cuts, for example due to absent time pressure, it is easier to memorize a detour through a green passage. Last, although urban density has been identified as important characteristic for salience and legibility of an environment, its role as a determinant remains unknown, as neither model indicated significance.

Gini coefficients can be used to identify dispersion and clustering of salient urban areas.

Results from the Gini coeficients demonstrate that it is more likely to encounter more local salient urban areas when moving across the historical city center of Amsterdam. Hence, routes across the historical city center are expected to be easier to memorize and structured in long-term memory.

The results provide empirically underpinned behavioural insights to improve management of urban tourism.

Based on the insights of activity patterns, the municipality can be advised to explore three measures. First, ensure that major routes between connected activity zones are well equipped for bicycle traffic of tourists and residents (slow speed/recreational paths and high speed/efficient paths), followed by allocating good wayfinding systems. Third, capacity issues concerning bicycle-parking places can be evaluated based on the identified activity zones while incorporating the expected growth of both commuters and tourist volumes.

1.7.2.3 Travel information

Discussing the possibilities and relevance of wayfinding styles for identification of potential navigation problems.

Both wayfinding styles can be used complementary as different processes influence them. However, two determinants (navigational preference to follow the bearing line and average daily distance travelled by car) have an ambiguous effect on both wayfinding styles. This could indicate a trade-off, because gathering and processing more spatial knowledge will ultimately require more orientation ability in order to process the knowledge into useable wayfinding strategies. The navigational preference to follow the bearing line is not beneficial when there is a low amount of spatial knowledge, as this does not encourage the acquisition of more spatial knowledge. If a satisfactory amount of spatial knowledge has been acquired using the bearing line as a navigational preference is useful to reduce the workload.

Discussion on insights considering how tourists can be better spatially and temporarily distributed.

To ensure local communities in outer areas also economically benefit from tourism, the possibilities and bicycle travel times should be better communicated and adapted to tourists standards. Many activity zones identified by k-means cluster analysis in the outer areas of Amsterdam are within a 45 minute bicycle trip (considering an average bicycle speed of 8 km/h) from the city centre, which is lower than the average duration of a bicycle trip of tourists that participated in the LUCY study, which is 48 minutes. Promotion of attractions and bicycle tours can be achieved through, for instance, tailored (seasonal) urban bicycle maps for tourism, including official parking places and day-itinerary suggestions to avoid the crowd, and stimulation of residents of socially deprived neighbourhoods and outer areas to organize local cycling tours during the summer period. Additionally, Pop-up events and new markets located within ~2.23 kilometres (i.e. the average Euclidean distance between activities) from an activity zone of the Top 15 are potential attractions to distribute tourists within the city while travelling by bicycle. Furthermore, it is important to ensure that a certain degree of variation exists within the salient region. Additionally, within a Top 15 activity zone visitors can be distributed to alternative attractions at walking distance (e.g. near Museum Square activity zone there are the less famous Amsterdam Art Station, Zuiderbad, Café Loetje and Wildschut). This measure will relieve the crowdedness at main attractions and increase the activity duration within activity zones, as visitors might visit both the main attraction and the secondary attraction. The results from the time patterns of tourists revealed a strong decrease in activity and movement intensities after 17:00. It is important to avoid big tourist flows during the morning and evening peak of residents. More major museums could explore longer opening hours (10 to 10), possibly for a reduced price, if tourists are travelling by bicycle. This could stimulate sustainable, healthy, and inclusive activity patterns, where tourists can visit more activity zones as alternative activity zones in outer areas can be visited at the start of the day. Further research is needed to assess the experience gained when exploring a city by bicycle and related benefits for residents as well as tourists.

1.8 Thesis Outline

Four research papers provide the background to discuss the main contributions in relation to the main and key research questions that are identified in this section. The first paper describes a theoretical framework of urban wayfinding styles and investigates the relation between two identified styles to travel behaviour, navigational preferences, urban environment, and socio-demographics. In corresponding chapter 2 we define the relevant dimensions that are required to charactise urban wayfinding styles and we unravel how they relate to daily travel behaviour. The second paper contains the relation between spatial analysis and urban cognition and is detailed in chapter 3, where a methodology is developed to use open spatial data for the identification of salient and legible urban areas (landmarks). A third paper unravels bicycle travel behaviour of tourists in Amsterdam based on an empirical GPS data set. The spatial and temporal characteristics of activity patterns are detailed in corresponding Chapter 4. The final paper describes their routing patterns and the role of spatial knowledge acquisition. The analysis and model results using the spatial boundaries of the route selection space can be found in Chapter 5.

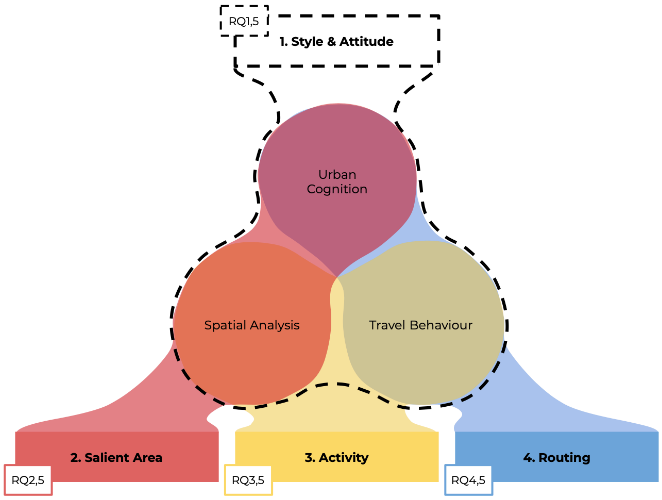


Figure 3. Thesis Outline.

1. Determinants of Urban Wayfinding Styles. We present a theoretical framework to relate wayfinding with travel behaviour, urban environment, and navigational preferences. We use travel diaries and questionaires to test several theories. (Theory = data).
2. Spatial Analytics to Identify Salient Areas. We present a methodology to characterisze salient and legible areas and demonstrate the potential of open urban data. (Data).
3. Activity Patterns of Tourists in Amsterdam ft GPS data from bicycles. We conducted an empirical study to characterize bicycle activity patterns of tourists. Large scale GPS data collection to retrieve insights on urban tourism behaviour and spatial dispersal. From the insights we construct new theories. (Data → insight → theory).
4. On the Relation between Learning the City and Routing. We Modelled the route choice ellipse dynamics based on detour ratio, maximum deviation, eccentricity and curcature. We use GEE models to unravel the structure behind the data, and investigate alignment with theories. (Theory = Model(data)).

The final chapter (6) of the thesis concludes with answering the main and key research question(s), and contains the reflection and recommendations for future work.

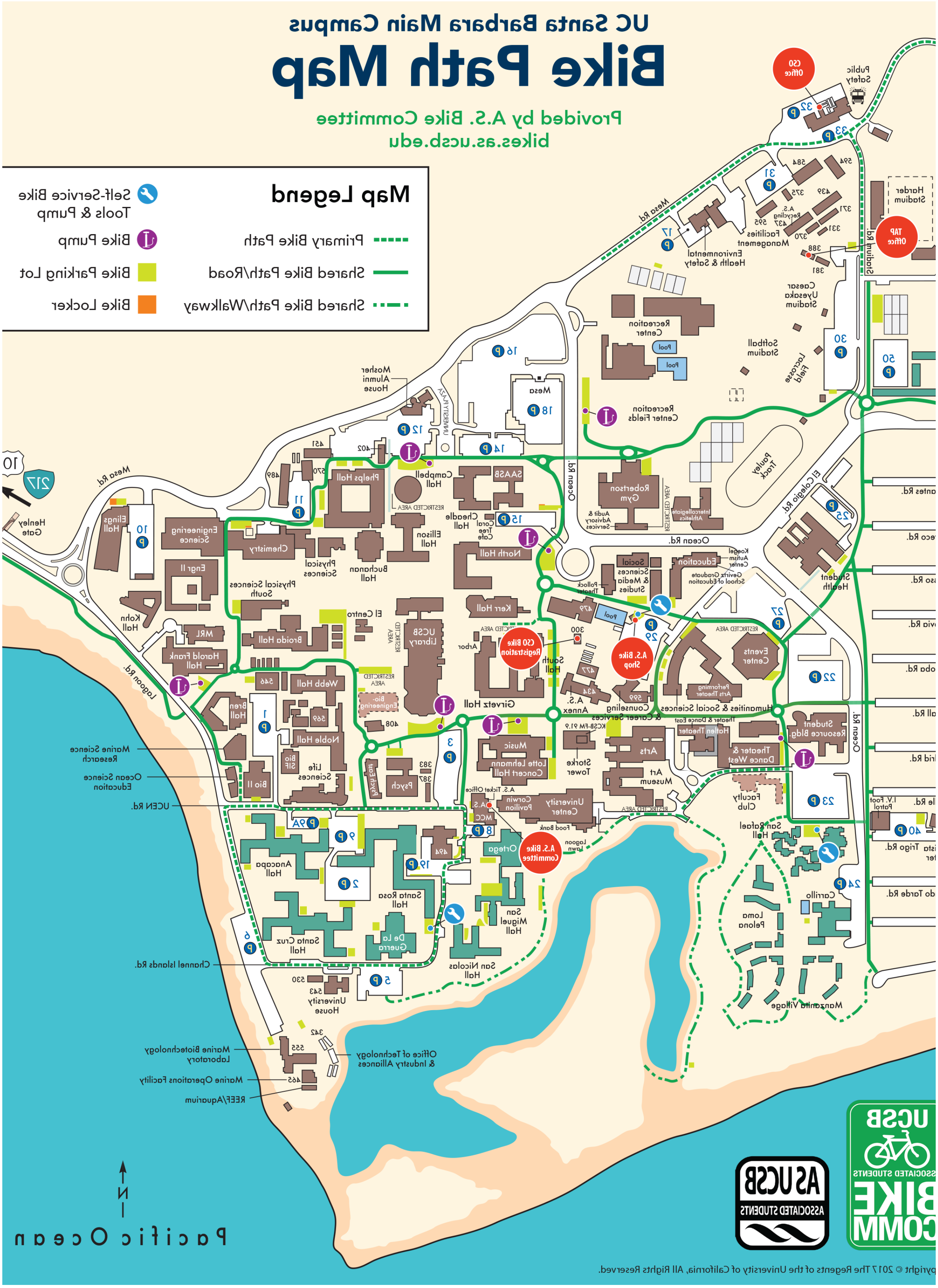
UC Santa Barbara Main Campus

Bike Path Map

Provided by A.S. Bike Committee

bikes.as.ucsb.edu

Copyright © 2017 The Regents of the University of California, All Rights Reserved.


UC Santa Barbara Main Campus
Bike Path Map
Provided by A.S. Bike Committee
bikes.as.ucsb.edu
Map Legend
Primary Bike Path
Shared Bike Path/Road
Shared Bike Path/Walkway
Self-Service Bike Tools & Pump
Bike Pump
Bike Parking Lot
Bike Locker
CSO Office
TAP Office
CSO Bike Registration
A.S. Bike Shop
A.S. Bike Committee
Public Safety
Harder Stadium
Facilities Management
Environmental Health & Safety
Caesar Uyesaka Stadium
Recreation Center
Softball Stadium
Lacrosse Field
Recreation Center Fields
Pauley Track
Mosher Alumni House
Mesa Rd.
Ocean Rd.
El Colegio Rd.
Robertson Gym
Intercollegiate Athletics
Student Health
Education
Social Sciences & Media Studies
Phelps Hall
Campbell Hall
SAASB
Cheadle Hall
Ellison Hall
North Hall
Buchanan Hall
Chemistry
Physical Sciences
Physical Sciences South
Engineering Science
Engr II
Elings Hall
Henley Gate
Kohn Hall
MRL
Broida Hall
Harold Frank Hall
UCSB Library
Arbor
Kerr Hall
South Hall
Events Center
Performing Arts Theater
Humanities & Social Sciences
Girvetz Hall
Counseling & Career Services
Webb Hall
Noble Hall
Life Sciences
Bio II
Psych
Music
Lotte Lehmann Concert Hall
Storke Tower
Arts
Art Museum
Theater & Dance West
Student Resource Bldg.
Faculty Club
University Center
Corwin Pavilion
Marine Science Research
Ocean Science Education
UCEN Rd.
Anacapa Hall
Santa Rosa Hall
Santa Cruz Hall
De La Guerra
Ortega
San Miguel Hall
San Nicolas Hall
San Rafael Hall
Carrillo
Loma Pelona
Manzanita Village
Channel Islands Rd.
University House
Marine Biotechnology Laboratory
Marine Operations Facility
REEF/Aquarium
Office of Technology & Industry Alliances
Lagoon Rd.
Pacific Ocean
N
AS UCSB ASSOCIATED STUDENTS
UCSB ASSOCIATED STUDENTS BIKE COMM
Copyright © 2017 The Regents of the University of California, All Rights Reserved.

2

Urban wayfinding behaviour

Everyday people find their way towards work, supermarkets, or unfamiliar places are explored for a social visit. Understanding how differences in urban wayfinding behaviour relate to daily travel patterns is important to describe route choice behaviour, identify potential navigation problems, design more legible cities, and provide comprehensible travel information. Therefore, the goal of this chapter is to jointly investigate the differences between urban wayfinding styles and the relations with socio-demographic, motility, urban environment, navigational preferences, and daily travel behaviour.

The main findings of the study are based on a sample of the Dutch population of 1.101 respondents that completed a three-day travel diary as part of the Mobility Panel Netherlands (MPN) and an additional cross-sectional survey designed to capture perceptions, attitudes, and wayfinding for active modes (PAW-AM). The five highlights are:

1. Contribution to theoretical insights of how urban wayfinding behaviour relates to daily travel patterns.
2. Execution of a large empirical analysis, to investigate the combined effects on two identified wayfinding styles using an exploratory factor analysis and Generalized Linear Models.
3. An objective and critical evaluation of the model results based on contingency tables and confusion matrices.
4. Discussion of the possibilities and relevance of wayfinding styles for route choice behaviour, provision of comprehensible travel information, design of legible cities, and identification of potential navigation problems.
5. Results provide evidence that predominantly different processes describe each wayfinding style.

This is an edited version of the following article:
Zomer, Schneider, Ton, Hoogendoorn-Lanser, Duives, Cats, & Hoogendoorn (2019). Determinants of urban wayfinding styles. *Travel Behaviour and Society*, *17*, 72-85.

2.1 Introduction

Each trip requires people to make various decisions before and during travelling. These decisions regard which modes and routes are to be used, and which activities will be performed where and when. Due to individual differences in preferences (e.g. minimize turns and thus choosing a simpler yet longer route) the urban experience differs, and as a consequence, the mental representation of the environment is likely to be different. In turn, these differences will influence future travel decisions resulting in different choice behaviour. Wayfinding behaviour is commonly defined as *the strategies that people use to decide how to move from one place to another* (Montello, 1995). It relates to the set of preferences, selection, and application of navigational strategies, the attitude towards travelling, and the ability to reach the intended destination. Differences in travel behaviour are expected to determine the extent to which wayfinding styles and navigational preferences are important to individuals.

Understanding how urban wayfinding behaviour relates to daily travel patterns is important to describe differences in route choice behaviour, identify potential navigation problems, design more legible cities, and provide comprehensible travel information. However, to date, relations between urban wayfinding styles and the complexity of daily travel behaviour, urban environment, and navigational preferences are largely unknown. Recent advances in cognition and travel behaviour research increased the understanding of the impact of socio-demographic factors on wayfinding and navigation behaviour of people through fMRI, (virtual reality) experiments, and questionnaires (Andreano and Cahill, 2009, Golledge et al., 1995, Prestopnik and Roskos-Ewoldsen, 2000). Commonly, these studies are conducted amongst small groups of undergraduates, using controlled experiments in small-scale environments primarily interested in the influence of gender and age (Maguire et al., 1999). Nonetheless, there are indications that active navigation (e.g. being the driver while driving or bicycling) relates to the ability to solve wayfinding tasks that require "route" and "map/survey" knowledge (Nori and Giusberti, 2006).

Wayfinding styles and navigational preferences in this chapter stem from a cross-sectional survey specially designed to capture perceptions, attitudes, and wayfinding for active modes (PAW-AM). A total of 1.101 respondents not only completed this survey, but also a 3-day travel diary, personal, and household survey as part of the longitudinal Mobility Panel Netherlands (MPN) Survey in 2016 (Hoogendoorn-Lanser et al., 2015). Wayfinding styles are based on the standardized self-report questionnaire of environmental spatial skills (SBSOD) originally developed and tested at the University of California-Santa Barbara (Hegarty et al., 2002). In explaining urban wayfinding behaviour based on literature on experimental studies, the variables of interest can be divided into four groups: *socio-demographic and mot*ility (e.g. gender, age, mode availability, and financial compensation), *urban environment* (e.g. urban density and perceived accessibility levels), *navigational preferen*ces (e.g. minimize number of turns and active navigation ratio), and *daily travel behaviour and patterns* (e.g. mobility portfolio, mobility cluster pattern). The objective of this chapter is to investigate how these determinants jointly relate to urban wayfinding styles. To this end, based on the SBSOD a Factor Analysis has been conducted to identify how many, and which, urban wayfinding styles exist. Generalized Linear Models (GLMs) are used to estimate to what extent various determinants affect two hypothesized urban wayfinding styles, in this chapter coined as Orientation Ability and Knowledge Gathering & Processing Ability.

The structure of the remainder of the chapter is as follows: the next section (Section 2.2) provides background on (urban) wayfinding behaviour. A description of the data and research approach is provided in Section 2.3. A Factor Analysis to derive the urban wayfinding styles from the self-reported preferences is provided in Section 2.4. The modelling results and relevance of the GLMs for the urban wayfinding styles are described in Section 2.5. We then synthesize the findings of this chapter in Section 2.6. This chapter finishes in Section 2.7 with a conclusion.

2.2 Literature background

This section provides an overview of wayfinding behaviour and determinants that have been found to impact wayfinding based on experimental studies. The remainder of this section first elaborates on the definition of wayfinding behaviour, followed by a synthesis of the main findings in relation to four categories of variables that are found in literature: socio-demographic and motility, urban environment, navigational preferences, and (daily) travel behaviour.

2.2.1 Urban wayfinding behaviour

Although strongly related, a distinction can be made between wayfinding, orientation, and navigational strategies. Wayfinding behaviour is typically associated with the *exploration* of the (possible) route(s) between an origin and a destination, given the urban network (Passini, 1980, Golledge, 2004). In this chapter, wayfinding behaviour encompasses two styles based on the attitudes towards spatial (network) knowledge and the orientation attitude. In theory, the combination of these wayfinding styles influences the boundary of the considered choices in daily travel behaviour, e.g. when deciding where to go, which travel mode to use, or which route to take. In turn, travel choices result in a specific urban experience that may stimulate to different wayfinding abilities and navigational preferences. Hence, we postulate that there might be a bi-directional relation between these two notions. A navigational strategy is more *goal-oriented* and is aimed at arriving at a destination with reference to a specific objective such as minimize travel time or distance (Hund and Minarik, 2006, Baldwin, 2009). Navigational strategies are considered as a preference that may be associated with different wayfinding styles. Spatial orientation is one of the wayfinding abilities, it is the *ability to identify and recall places from different physical positions and graphical representations* (Gärling et al., 1986). This chapter investigates *wayfinding styles* based on the standardized self-report questionnaire of environmental spatial skills (SBSOD) relating to attitudes towards spatial knowledge acquisition (exploration), orientation within, and mental representation of, the environment, anxiety, and usage of route information (Hegarty et al., 2002).

Most wayfinding studies have investigated to what extent two hypothesized wayfinding styles, *route-based* and *map/survey based*, can describe how individuals find their way (Foo et al., 2005, Hund and Minarik, 2006, Xia et al., 2008, Carlson et al., 2010). These studies depict Route-based wayfinding as more or less an egocentric orientation style (with memorized sequences of local views) along a route. Consequently, specific decisions and actions are associated with landmarks, intersections, and sights. Whereas, map-based wayfinding is used when orientation is considered to be allocentric and/or coordinated. In the latter style, the developed mental map includes spatial relations and distances between important urban elements. Especially in urban environments the ability to orient and memorize the current position, and to construct a mental representation are crucial, as moving through a city requires one to integrate the sequence of views that change with one's movement in the environment (Hegarty et al., 2006). The likeability that identified styles can be disassociated at different scales of space has been investigated using factor analyses (Hegarty et al., 2006). To date, the extent to which these wayfinding styles relate to daily travel behaviour (e.g. activity and route choice behaviour) in the urban environment remains unclear.

To investigate the extent in which daily travel behaviour (e.g. activities, mode use and travel distance) actually explains urban wayfinding behaviour this chapter builds on top of the majority of wayfinding studies by proposing a theoretical framework inspired by literature on wayfinding behaviour (Stea and Blaut, 1973, Siegel and White, 1975, Golledge and Gärling, 2001). An Exploratory Factor Analysis is used to reduce the dimension of wayfinding styles from the SBSOD into components of mutually exclusive wayfinding styles. Each factor component can be divided into three levels of a wayfinding style; lower than average, average, and higher than average.

2.2.2 Determinants of wayfinding behaviour

The aim of this section is to describe behavioural insights related to preferences, attitudes and urban wayfinding behaviour reported in literature. Wayfinding behaviour based on self-report questionnaires, such as the SBSOD, is widely and extensively investigated in various fields, ranging from neurosciences and psychology to anthropology. To compare our findings with existing studies, the most relevant studies have been selected based on three requirements: 1) Reference to the original SBSOD questionnaire, 2) New data collection efforts, and 3) The analysis contains at least one determinant related to the urban environment, navigational preferences, and (daily) travel behaviour. This section is limited to the identification of some general trends in research methodology. The remainder of the section describes the key determinants that have been investigated in relation to wayfinding and travel behaviour.

2.2.2.1 Trends in research methodologies to study wayfinding behaviour using the SBSOD

Research into wayfinding behaviour dates back to the beginning of the 20th century. However, most determinants have been systematically investigated from 1990 onward. The majority of the studies conducted in the past two decades are controlled experiments, where a small number of participants are asked to complete a specific order of predefined tasks in a delineated spatial environment. These studies required time-intensive data collection efforts per participant (e.g. fMRI, 1-on-1 shadowing of movements, VR studio). As a result, the median sample size is 32 participants, and merely 8 studies have a sample size of more than 100 participants. Consequently, there are not many studies that aimed to relate multiple aspects with wayfinding behaviour and often only the influences of gender and age have been investigated. Regardless of the number of participants, most studies apply (multivariate) analysis of variance ((M)ANOVA) with a limited number of determinants and/or Structural Equation Models (SEM).

Although most studies included a(n adapted) version of the SBSOD, the processing of the responses varied from verification of non-significant differences within clusters in the sample to an average score for all questions, or a Factor Analysis. Also, sense-of-orientation (as a wayfinding style) has been used both as explanatory as well as dependent variable. As a consequence, discrepancy exists in results for many determinants. For example, there is no unified consensus on the relation with gender, the most researched determinant. Interestingly, also within research groups, findings and conclusions vary, which leads to critical theoretical reflection studies (Shelton et al., 2013, Piccardi et al., 2011).

Several research gaps can be identified related to the urban environment, such as the extent to which the larger metropolitan urban environment, where daily travel behaviour takes place, affects urban wayfinding behaviour. Also, while urban density has been identified as an important characteristic for salience and legibility of an environment, its role as a determinant remains unknown (Brunyé et al., 2010, Hölscher et al., 2011, Emo, 2012, Chrastil and Warren, 2014, Li and Klippel, 2016).

Navigational preferences are analysed using verbalized reports of respondents while walking or driving along a predefined route (Kato and Takeuchi, 2003, Hölscher et al., 2011, Arnold et al., 2013, Meilinger et al., 2014, Weisberg and Newcombe, 2016). A strong focus on wayfinding efficiency in many studies leads to a rather subjective classification such as good and bad orienteers while refraining from these definitions allows a deeper understanding of the versatility exercised by individuals (Shelton et al., 2013). However, due to the lack of extensive studies, the low number of participants, and differences in operationalization of both wayfinding and navigational preferences, it remains insofar unclear how daily travel behaviour in urban environments influences which navigational strategies are preferred, and if differences exist in relation with wayfinding behaviour.

Furthermore, very few experimental studies are positioned within the travel behaviour research field. However, relations between wayfinding and travel behaviour have been investigated in numerous studies at the operational (route) level, leaving a research gap on the influences on tactical and strategic levels relating to mode, route and activity choices, but also on the average daily travel behaviour. Most of these studies are goal-directed while there is empirical evidence to suggest that specific tasks influence wayfinding and search behaviour (Emo, 2012). Some studies investigated the relation between wayfinding and daily travel behaviour, but either based on a homogeneous and often very specific sample (elderly or children (Turano et al., 2009, Phillips, 2013, Taillade et al., 2016), students (Kato and Takeuchi, 2003, Hegarty et al., 2006, Ishikawa and Kiyomoto, 2008), Yucatec Maya farmers (Cashdan et al., 2016), limited to one travel mode (pedestrian (Li, 2006, Arnold et al., 2013, Giannopoulos et al., 2014), taxi/car (Turano et al., 2009, Han and Becker, 2014)), or a qualitative assessment of verbalized responses (Phillips et al., 2013). Hence, it remains unclear how wayfinding styles translate to urban navigation and daily travel behaviour in practice, and how individual characteristics potentially mediate differences, also acknowledged by Shelton et al. (2013). Moreover, to increase realism and establish a more extensive framework there is a need for a more heterogeneous and diverse sample (e.g. not only walking behaviour, students, children, elderly, or women) and move beyond extremely unrealistic environments (e.g. simplified VR mazes) into common wayfinding situations.

2.2.2.2 Wayfinding determinants in relation to travel behaviour in literature

This section elaborates on the main and most striking determinants that have been investigated in relation to wayfinding and travel behaviour. To summarise, the experimental studies that included travel behaviour characteristics are conducted at operational route level, leaving a research gap on the influences of tactical and strategic levels relating to mode, route, and activity choices, but also on the average daily travel behaviour.

Most studies were performed in a North American, European (UK and Germany), or Japanese context, where the majority of the findings lean towards the hypothesis that men are better at orientation and navigational tasks, while women have enhanced knowledge gathering, memory and processing ability. Results indicate that different cognitive strategies are used; men rely more on Euclidean distance and direction, whereas women prefer to find their way based on salient landmarks (Schmitz, 1997, Kimura, 2000, Waller, 2000, Lawton and Kallai, 2002). The difference can become more apparent depending on the environment (Silverman et al., 2000, Malinowski and Gillespie, 2001, Saucier et al., 2002, Andreano and Cahill, 2009). Research focusing on brain activity using fMRI studies investigated socio-demographic differences during tasked viewpoint resemblance based on photographs (Epstein et al., 2005), wayfinding in a museum (Janzen et al., 2008), and relations between memory engagement, navigational learning strategies, and percentage of finding destinations using short-cuts (Furman et al., 2014).

With aging societies, there is increasing research interest in wayfinding abilities and difficulties among the elderly. Turano et al. (2009) investigated mobility levels, described by the visit frequency to neighbouring areas, of elderly by car in Maryland (USA). Phillips, 2013, Phillips et al., 2013 explored the influences of landmarks and complexity of street layout, familiarity, various navigational preferences, and trip purpose on how the elderly experience urban environments. Also, differences at operational travel behaviour have been investigated, such as the frequency of stops and detours due to decreased orientation (Taillade et al., 2016).

In travel behaviour research it is common to also include household characteristics, such as the number of children, as this poses limitations to the flexibility and induces certain type of activities. This is sometimes also be regarded as motility, the potential and ability to move (Lucas, 2012). Slightly different from individual or household characteristics, motility relates more to availability of, and accessibility to transport modes, monetary compensation that may affect the affordability, and childhood experiences that may affect the development of initial wayfinding behaviour (experienced motility). A unique study on the Yucatec Maya farmer community in a rural and remote area showed that mobility, based on the number and frequency of visits to various sights in the region, has a direct relation with an interaction effect of gender and marital status, which would explain gender differences in self-reported wayfinding styles. During childhood there are no significant differences in mobility patterns between boys and girls, only once married Yucatec men start to travel to more distant areas, while Yucatec women stay more frequently at home (Cashdan et al., 2016). Slightly different are studies using the NASA Task Load Index (NASA TLX) to assess subjective cognitive workload due to difficulties with navigating using (innovative) travel information applications (e.g. Baldwin, 2009, Rehrl et al., 2012). Provision of more information will only benefit those with suitable wayfinding abilities and aligned navigational preferences to successfully process the information.

Nearly two-thirds of the studies included at least one determinant describing travel behaviour. Travel behaviour has commonly been described at the operational or route level where participants either were requested to follow a predetermined route, or walk to a predetermined location within delineated environments (Hölscher et al., 2011, Ishikawa and Nakamura, 2012, Emo, 2012, Chrastil and Warren, 2014, Taillade et al., 2016, Li and Klippel, 2016). Findings from these studies consistently indicate that travel distance and time have a negative relation with the wayfinding score obtained in the SBSOD. Also, local and global salience of a location influences how easy it is to find a destination, which affects the travel distance and time. Participants with lower wayfinding scores also make more errors while finding their way and show higher workloads when required to use travel information services.

Several studies investigated the relation with wayfinding behaviour and more common daily travel characteristics, mainly by asking the frequency of visits to certain neighbourhoods or locations (Nori and Giusberti, 2006, Turano et al., 2009, Piccardi et al., 2011, Phillips, 2013, Cashdan et al., 2016). Travel behaviour at route level is usually investigated in the form of a description of the differences between chosen alternatives in terms of distance, turns, or most traversed intersections (Hölscher et al., 2011, Furman et al., 2014).

Based on past findings, it can be expected that respondents with a better sense of orientation choose routes with shorter travel distance and time, but not necessarily higher travel speed. This requires flexible navigational preferences as the structure and layout of each urban environment demands different abilities. To the authors' knowledge, the relation between wayfinding behaviour and real-life daily travel behaviour has not been quantified except for specific target groups.

2.3 Research approach & methodology

To investigate to what extent urban wayfinding behaviour can be described from a holistic perspective, by jointly including urban environment, navigational preferences, and daily travel behaviour, we have enriched the longitudinal Mobility Panel Netherlands (MPN) with a cross-sectional survey in 2016. This additional survey (PAW-AM) is designed to capture perceptions, attitudes, and wayfinding for active modes, and included a Dutch version of the standardized self-report questionnaire called the Santa Barbara Sense of Direction (SBSOD). Upon aggregation of both questionnaires (MPN 2016 and PAW-AM) to the individual level, various data processing techniques have been used to derive determinants of interest, such as a latent class cluster analysis (LCCA) to capture daily mobility patterns, instead of separate determinants describing daily travel behaviour characteristics (Ton et al., 2019).

To operationalize wayfinding styles, an Exploratory Factor Analysis based on the standardized self-report SBSOD is performed. Generalized Linear Models (GLM) are used to identify how differences in wayfinding styles can be explained by socio-demographic, (perceived) urban environment, navigational preferences, and daily travel behaviour. These research steps are visualized in Figure 4 and further detailed in the remainder of this section.

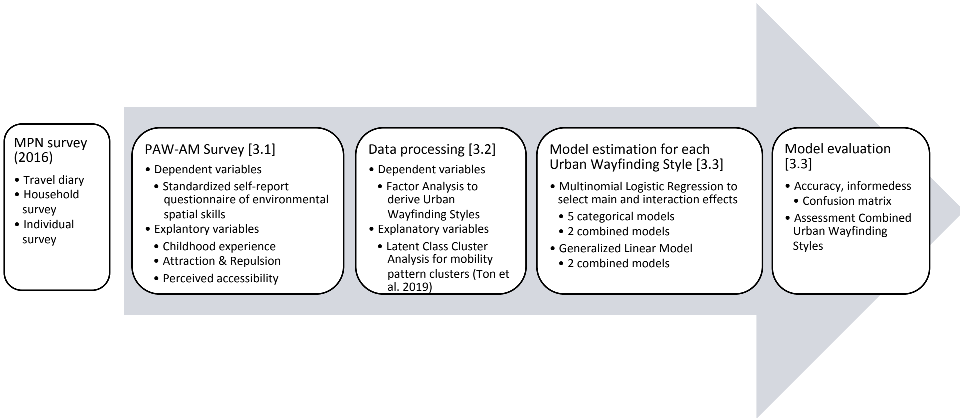


Figure 4. Research steps.

2.3.1 *Data on urban wayfinding*

The data used in this chapter stem from Dutch citizens that have completed a three-day travel diary, and personal and household surveys as part of the Mobility Panel Netherlands (MPN) in 2016 (Hoogendoorn-Lanser et al., 2015). The travel diary was computer-based and designed to provide information on activity and trip level. The travel diary provides different insights than commonly measured at route level or regional level travel behaviour in relation to wayfinding behaviour.

Enriching the MPN with cross-sectional special issues is important to better describe the underlying behavioural dynamics. To enhance the explanatory power of the MPN regarding pedestrian and cyclist mobility choices, the PAW-AM survey was designed. This survey featured among other things social norms and mode choice habits. One section of the PAW-AM survey was related to urban wayfinding behaviour and navigational styles. The PAW-AM survey has been designed to reduce respondents' fatigue, as such; half of the respondents of the PAW-AM received questions related to wayfinding behaviour, while the other half received questions related to social norms. Thus, the presented analyses and models stem from 1.101 respondents that completed a 3-day travel diary, personal and household survey, and the PAW-AM survey focused on wayfinding behaviour in order to investigate which determinants, and to what extent, relate to urban wayfinding styles.

Urban wayfinding variables in the MPN and PAW-AM. Urban wayfinding behaviour is investigated based on the SBSOD (Hegarty et al., 2002). The focus of this questionnaire is on the attitudes towards spatial knowledge acquisition (exploration), orientation within an environment, mental representation of the environment, anxiety, and usage of route information. All respondents are asked to indicate how much a statement reflects their behaviour, ability, or attitude at 5-point Likert-scale (1: strongly disagree and 5: strongly agree). All questions were translated to Dutch, and approximately half of the questions are stated positively, and half negatively. In total 23 statements have been used, two examples are: "*I easily get lost in a new city*" and "*I enjoy reading maps*".

The explanatory variables in this chapter are visualized in Figure 5 and further detailed in the remainder of this section. There are four independent variable categories: socio-demographic and motility, urban environment, navigational preferences, and daily travel behaviour.

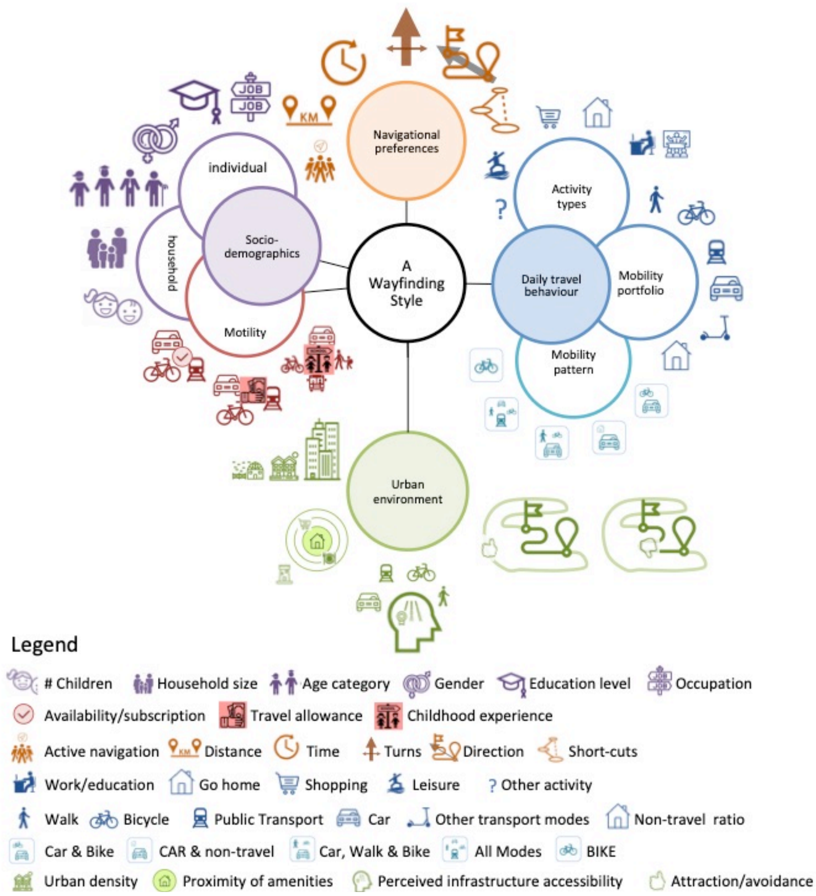


Figure 5. Conceptual framework of relations with urban wayfinding behaviour.

Individual and household characteristics such as gender, age, occupation, education, household size, and number of children are derived from the individual and household questionnaires of the MPN survey. Also, *motility* indicators of ownership of a car, bicycle and/or a transport subscription and eligibility to any form of compensation for a certain mode by the employer or special discount for low-income households are derived from the MPN survey. An additional subgroup related to motility has been included in the PAW-AM survey, namely, whether people during their childhood experienced travelling to school by foot, bike, public transport, or were driven by car. This metric is intended to provide an indication of the size of travel environment at the age when people are likely to start developing their wayfinding skills. In line with the cultural-behavioural-brain (CBB) loop model (Han and Ma, 2015), the underlying hypothesis is that these “first” experiences may influence today’s attitude and perception towards travelling

Variables related to the physical urban environment are derived from the MPN survey, where a high-level indication is available regarding the urban density in the region (rural, urban, or highly urbanized region). The PAW-AM survey focused on the perception of the *urban environment*, and included statements such as "*in my neighbourhood there are shops/restaurants/old buildings within walk or bicycle distance*", and "*the infrastructure in my neighbourhood is walk/bicycle/public transport/car friendly*". In the first section of the PAW-AM survey participants are asked to identify for which trip purpose they used the bicycle most often. In the section featuring navigation styles, respondents were asked to identify from a list of 26 urban elements which urban elements they would avoid on their way to this activity. The English translation of this question is: "*I am willing to* ***make a detour*** *when I* ***cycle*** *to* ***[personalized trip purpose]*** *if I can* ***avoid …***". The urban elements were very diverse, ranging from "crowded bicycle paths" to "streets with townhouses", and "areas where many traffic accidents happen". The chosen elements were classified with the label 'negative'. From the list of urban elements remaining after their first selection, respondents were asked to identify which urban elements would attract them, which were accordingly labelled 'positive'. All items that have not been indicated as repellent or attractive are classified as 'neutral'.

Navigational preferences relate to the decision-making strategies to choose or follow a specific route. The PAW-AM survey includes 5-point Likert scale questions regarding five navigational preferences minimizing (i) travel distance, (ii) travel time, (iii) number of turns, or (iv) following the direction (bearing) towards the destination, and (v) taking short-cuts. Active navigation ratio has been derived from the 3-day-travel diary and depicts how often a respondent has actually been in charge of wayfinding for a certain trip, for example as the driver of a car, and any bicycle or walking trip. In this chapter, the average active navigation ratio is 0,70 (70%) with a standard deviation of 0,42. The related hypothesis is that individuals with a higher active navigation ratio have more advanced wayfinding abilities, as one relies more often on his or her abilities.

All *travel behaviour* characteristics are aggregated based on the average behaviour reported during weekdays that needed to be reported in the 3-day travel diary as part of the MPN in 2016. Note that respondent A with 3 trips on Monday, 1 trip on Tuesday, and 1 trip on Wednesday yields an average daily number of trips of 1,67, with a non-travel ratio of 0. Respondent B with 2 trips on Sunday (weekend), 0 trips on Monday, and 3 trips on Tuesday yield an average daily number of trips of 1,5, with a non-travel ratio of 0,5, because only two weekdays needed to be reported. The non-travel ratio is important because it shows the share of active days. Trip purpose (activity type) has been classified into 5 categories: (i) going to work/school, (ii) going back home, (iii) doing (grocery) shopping, (iv) performing a leisure activity (sports, restaurant), and (v) other activities. Commonly, shopping, sightseeing, social visits to family and friends require people to travel through, or to, unfamiliar environments (Phillips, 2013).

2.3.2 Derivation of urban wayfinding styles

Previous work described in the Literature Background (Section 2.2) established that the SBSOD has a high degree of test–retest reliability. SBSOD scores are able to predict performance on experimental tests that require subjects to update one's location and orientation in space. In order to reduce the dimensionality of 23 questions of the self-report questionnaire, an Exploratory Factor Analysis is used to derive urban wayfinding styles. Prior to the factor analysis, negatively stated questions were reversed to derive a positive relation with each component. To minimize multicollinearity effects and to identify the underlying dimensions of urban wayfinding styles, principal component extraction and varimax rotation have been applied. The resulting components of the factor analysis constitute a set of latent variables that describe environmental spatial skills and urban wayfinding behaviour. Categorization of the latent variables to three levels with cut-off values at −0,5 and 0,5 transforms the latent variables to three wayfinding styles per component: lower than average (-1), average (0), and higher than average (1). The results of the factor analysis and components are described in 2.4.1.

2.3.3 Model estimation of urban wayfinding styles

The goal of this chapter is to answer the research question "To what extent can differences in the urban wayfinding styles coined as 'Orientation Ability' and 'Knowledge Gathering & Processing Ability' be explained by a comprehensive model including the relations with socio-demographic and motility, urban environment, navigational preferences, and travel behaviour?" This can be investigated using different statistical models, including Generalized Linear Models (GLM) (Nelder and Wedderburn, 1972, Diggle et al., 1994, Cox et al., 2013) and multinomial logistic regression. One of the major pitfalls of multinomial logistic regression is the reduction in degrees of freedom when many parameters are included. Different from regression models, GLM assumes that there is no clustering of the data and thus responses of all respondents are mutually independent.

GLMs consist of three components: the systematic linear prediction, a multinomial random component, and a Logit link function (Nelder and Wedderburn, 1972, Diggle et al., 1994, Cox et al., 2013). The first component is similar to OLS regression as it describes the linear relation between a function of the expected dependent variable $Y = g^{-1}(\eta)$, and the explanatory variables in the model,

$$Y = g^{-1}\left(b_0 + \sum_{i \in j} b_i X_i\right) \tag{1}$$

Where,
Y = measured ability level of a wayfinding style (low, average, or high)
g = Logit link function, transforms the predicted value of the dependent variable (η) to a new form that has a linear relationship with Y
b_0 = intercept
b_i = estimated weight coefficient for a given explanatory variable i
X_i = explanatory variables (age category, average daily distance by car ed.)

The link function allows for non-linear relations between explanatory variables and the predicted outcome. Applying the parallel regression assumption, the link function transforms the expected value of Y to a new form that has a linear combination of the explanatory variables, ordered from high to low with respect to the highest level. The model output includes b-coefficients to represent the average effect across the entire population of a change in X on the probability of the urban wayfinding style condition. For example, a 1-unit increase in an explanatory variable (i.e. age category) corresponds to a b_{age}-unit increase in the Logit of the expected value of "high Orientation Ability" versus lower conditions, holding all other variables in the model constant.

Before using Generalized Linear Models (GLM), a systematic approach of multiple multinomial logistic regression analyses has been used to provide a first selection of interaction effects and control for correlations between determinants in the holistic (combined) models. This systematic approach consists of four steps. To start, five categorical models are estimated for each of the variable categories (socio-demographic, motility, urban environment, navigational preferences, and travel behaviour). To derive a categorical model, first only main effects are included in a logistic regression using a forward stepwise method. The probability to include (resp. exclude) variables is 0,05 (resp. 0,10) with a maximum number of stepped effects of 40. Secondly, 2-way interaction effects are included, while insignificant main effects from step 1 are excluded. The outcome is a set of "primary determinants". Thirdly, all "primary determinants" are excluded in order to find "secondary determinants", under the premise that secondary determinants should have been significant as main effect in step 1. Fourthly, when all primary and secondary determinants of interest for all five variable categories are known, a last logistic regression is used to derive a combined model. Accordingly, for the combined model also interaction effects are included between categories, but only of variables that appeared significant at step 3 and 4. Given that relations between urban wayfinding ability and travel behaviour are of special interest, significant variables related to walking, bicycling, public transport, or car travel behaviour are always included in the combined model.

All determinants included in the two combined logistic regression models for Orientation Ability and Knowledge Gathering & Processing Ability are included in the estimation of two GLMs. To test the model fit, generally a Type I analysis is recommended when main effects are specified before first-order interaction effects. A model with an insignificant value implies that the related effect is not different from 0 if only the preceding effects are included. Therefore, next to a Type I test, also the Type III test is used to determine whether an effect is significantly different from 0 containing all modelled effects. Finally, model evaluation is performed using a confusion matrix to derive, amongst others, Accuracy and Informedness.

2.4 Theoretical framework & descriptive results

This section starts with the results from the Exploratory Factor Analysis and derivation of two urban wayfinding styles (2.4.1). Then, the preliminary results for daily travel behaviour based on travel diaries are described (2.4.2).

2.4.1 Factor analysis to derive wayfinding styles

After studying the Factor Analysis results for 2 to 5 components, it is concluded that the most applicable (consistent) number of components is two including 19 out of 23 questions. The results are depicted in Table 2.

Table 2. Rotated component matrix.

	1: Orientation Ability	2: Knowledge Gathering & Processing Ability
1. Sense of orientation[a]	0,75	
2. Ability to find the way in an unfamiliar city[a]	0,74	
3. Ability to understand route directions[a]	0,73	
4. Memorize a route after following it once	0,68	
5. Memorize a route as a passenger in a car[a]	0,68	
6. Ability to give route directions	0,67	0,44
7. Active navigation for longer journeys[a]	0,65	
8. Attitude to give route directions[a]	0,61	
9. Perception of distances	0,59	0,47
10. Perception of mental map[a]	0,57	
11. Ability to recall places[a]	0,54	
12. Attitude to read maps	0,54	0,48
13. Coordinated perception of environment (NSEW)		0,61
14. Exploration attitude to find new routes		0,68
15. Regularly choose new routes		0,76

[a] scored in reverse order as the survey question was negatively phrased.

The two components are coined Orientation Ability (attitude and basic skills to be able to orient and navigate effectively in an urban environment) and Knowledge Gathering & Processing Ability (attitude and preferences to extend knowledge about the environment, e.g. explore cities and take new routes). With this clustering, 50% of the total variance is explained. The KMO value is 0,940 and the Bartlett test indicates significance ($p < 0{,}001$). Each wayfinding style relies largely on unique variables, while also three common variables exist: ability to give route directions, perception of distances, and attitude to read maps. These results advocate that Orientation Ability and Knowledge Gathering & Processing Ability are partially dissociated.

Based on the resulting components a theoretical framework is proposed inspired by literature on wayfinding (Figure 6 A). In this framework *Orientation Ability* is the latent variable that captures three basic types of spatial orientation: egocentric, allocentric (fixed-point), and map-based (coordinated) orientation and navigation (Stea and Blaut, 1973). Similarly, *Knowledge Gathering & Processing Ability* is the latent variable for three basic types of spatial knowledge that can be acquired: declarative knowledge of landmarks, procedural route (network), and relational survey (map) knowledge (Siegel and White, 1975, Golledge and Gärling, 2001). However, the classification based on literature is not mutually exclusive, i.e. one can simultaneously rely on egocentric and fixed-point orientation. The factor analysis is used to transform the latent variables into unique wayfinding styles as it reduces the dimension to 2 components (Figure 6 B). Each component can be divided into three levels of a wayfinding style; lower than average (−1), average (0), higher than average (1). Using Generalized Linear Models (GLMs) the wayfinding styles can be investigated through the relationships with discrete and continuous variables related to socio-demographic, motility, urban environment, navigational preferences, and daily travel behaviour (Figure 6 C). The number at the centre of each box depicts the number of respondents diagnosed with each of the styles.

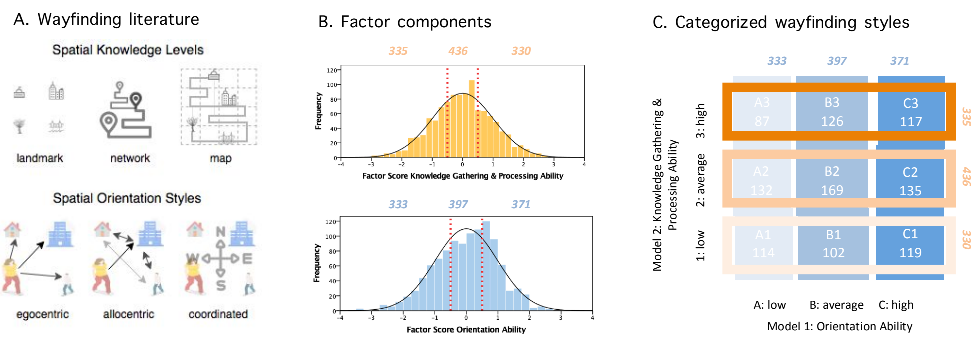


Figure 6 A-C. Theoretical framework (A-B) for operationalization of urban wayfinding styles and results (C).

2.4.2 *Daily travel behaviour*

Regarding daily travel behaviour in The Netherlands, almost 45% of the respondents did perform at least one bicycle trip during the three-day travel diary period, but every respondent in this chapter did make a bicycle trip in the past 6 months. From the 45% of respondents that used the bicycle during the travel diary period, 60% biked on average up to 6.0 km on a day. Furthermore, assuming an average cycling speed of 15 km/h, the results in Figure 7 suggest that 40% of the respondents are on average 24 to 46 min active on their bicycle on a daily basis. Moreover, approximately 37% of the respondents do not include any trips by car in their daily travel behaviour during the three-day travel diary period. From the respondents where the car is part of their daily travel pattern 20% travel only for short distances (not more than 9.0 km a day) Figure 7.

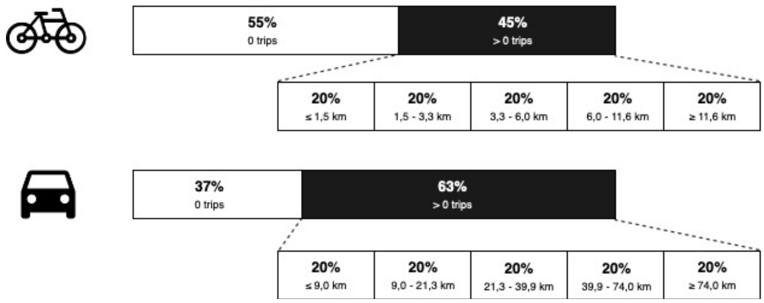


Figure 7. Mobility portfolio: average daily trips & travelled distance reported during travel diary period.

For this chapter, it is of interest to investigate to what extent *activity patterns* and *mobility portfolio* (frequency and distance of modal trips) relate to wayfinding styles. Furthermore, a *daily mobility pattern* is assumed to capture travel behaviour more realistically than individual mobility portfolio per travel mode, because consists of the combined intensity of all modes used. The MPN travel diaries are used to derive five mobility pattern typologies using a latent class cluster analysis. For more information on the technique used to derive the pattern typologies, the reader is referred to (Ton et al., 2019). For the estimation of the latent clusters, the average daily number of trips by car, public transport, bicycle, foot, other modes, and the share of non-travel days during the week have been used as input variables. The active covariates are urbanization level, occupation, and number of household members. The data concerning the complexity of daily mobility patterns includes stems from a larger set of 2.425 respondents.

2.5 Results

This section continues with the outcome of the Generalized Linear Models (GLM) for both urban wayfinding styles (2.5.1). After the explanation of model results, both models are evaluated based on confusion matrices (2.5.2).

2.5.1 Model estimation of urban wayfinding styles

Prior to the Generalized Linear Model (GLM) estimation, several multinomial logistic regression analyses have been performed to identify primary and secondary variables of interest for each variable category (socio-demographic, urban environment, navigational preferences, and daily travel behaviour). An overview of significant determinants based on the combined models with an urban wayfinding style can be found in Table 3. Three socio-demographic variables yield significant relations in one or both Generalized Linear Models (GLMs): gender, age, and education level. Regarding the urban environment, four determinants have been included related to attractiveness of urban elements, namely familiar streets, unfamiliar streets, greenopy, and rivers and lakes. Also, the perceived accessibility of the bicycle infrastructure in one's neighbourhood yielded significant differences in Orientation Ability. Note that merely 1–2% of the respondents avoid familiar streets, greenopy, or have a negative perception of the accessibility of the bicycle infrastructure. Therefore, it is likely that these variable levels will not yield significant results. Additionally, three determinants describing various facets of navigational preferences are found to be significant: preference to minimize turns, take short-cuts, and follow the bearing line (direction towards the destination). The latter four determinants were originally measured at a 5-point Likert scale. Active navigation ratio (0,0–1,0), average daily distance travelled by bike and car, and the average number of trips by car are included as continuous variables. Trip purpose and daily mobility patterns did not yield any significant relations with either wayfinding styles.

Table 3. Overview of determinants.

Discrete Variable Information	Levels	Frequency		Percentage	
gender	male	488		44%	
	female	613		56%	
education level	university degree	403		37%	
	vocational education	420		38%	
	basic education	278		25%	
age	>65 years	183		17%	
	30-65 years	504		46%	
	19-29 years	336		31%	
	12-18 years	78		7%	
perceived bicycle accessibility	good	959		87%	
	normal	126		11%	
	bad	16		2%	
familiar streets	attracted	261		24%	
	neutral	831		76%	
	repelled	9		1%	
unfamiliar streets	attracted	77		7%	
	neutral	917		83%	
	repelled	107		10%	
greenopy	attracted	441		40%	
	neutral	647		59%	
	repelled	13		1%	
rivers & lakes	attracted	16		2%	
	neutral	1068		84%	
	repelled	50		5%	
navigational preference minimize turns	(strongly) positive	275		25%	
	neutral	424		39%	
	(strongly) negative	402		37%	
navigational preference to take short-cuts	(strongly) positive	279		25%	
	neutral	492		45%	
	(strongly) negative	330		30%	
navigational preference follow bearing line	(strongly) positive	660		60%	
	neutral	360		33%	
	(strongly) negative	81		7%	
Continuous Variable Information		**Min.**	**Max.**	**Mean**	**Std. Dev.**
active navigation ratio		0	1	0,70	0,42
average daily distance by bike [km]		0	50,70	3,14	6,07
average number daily car trips		0	11	1,52	1,60
average daily distance by car [km]		0	454	30	49

2.5.1.1 Urban wayfinding style I: Orientation Ability

The first GLM model is used to estimate the latent ability for spatial orientation. The variables with the highest factor loadings are: sense of orientation, ability to find the way in an unfamiliar city, and ability to understand route directions. A total of 11 parameters are included (Table 4), the main effects of the GLM are gender (female or male), age (teenagers, young adults, middle-aged adults, or (young) seniors), perceived accessibility of bicycle infrastructure in the neighbourhood (bad, neutral, or good), preference to make detours via familiar streets and rivers or lakes (detour due to attraction or repulsion, or a neutral attitude), and the navigational preference to minimize turns and follow the bearing line towards the destination (disagree, neutral, or agree). Also four interaction effects yield significant results: gender and self-reported average daily distance travelled by bicycle (continuous measurement scale 0,0–50,7 km), self-reported average daily distance travelled by bicycle and preference to minimize turns, self-reported average daily distance travelled by car (continuous measurement scale 0,0–454,0 km) and preference to minimize turns, and self-reported average daily number of trips travelled by car (discrete measurement scale 0–11) and active navigation ratio (continuous measurement scale 0,0–1,0).

Table 4. Model results.

		Orientation Ability		**Knowledge Gathering & Processing Ability**	
Number of parameters (df)		11 (22)		6 (23)	
AIC | BIC		2.153 | 2.273		2.081 | 2.206	
Parameters		**B**	**Std. Error**	**B**	**Std. Error**
Threshold: low score		**-1,39*****	0,29	**-3,20*****	0,41
Threshold: medium score		0,42	0,30	**-1,27*****	0,39
Determinants	**Level/Scale**				
gender *[ref: male]*	female	**-1,36*****	0,14	**-0,68*****	0,12
age	12-18 years	**-0,82****	0,28	**-1,18*****	0,28
[ref: older than 65]	19-29 years	0,23	0,19	**-1,08*****	0,19
	30-65 years	0,39	0,17	**-0,80*****	0,18
perceived bicycle accessibility	bad	-0,39	0,37		
[ref: good]	normal	**-0,49****	0,18		
familiar streets	avoid	-0,75	0,55		
[ref: attracted]	neutral	**0,41****	0,15		
rivers & lakes	avoid	0,22	0,33		
[ref: attracted]	neutral	**-0,48***	0,18		
preference to minimize turns	disagree	**1,02*****	0,22		
[ref: agree]	neutral	**0,46***	0,20		
preference to follow bearing line	disagree	-0,38	0,25	0,18	0,30
[ref: agree]	neutral	**-0,39****	0,13	**0,29****	0,13
average daily distance by car	kilometre			**0,01*****	0,00

Interactions	Level/Scale	B	Std. Error	B	Std. Error
gender * bicycle distance	female	-0,02	0,02		
	male	**-0,08*** **	0,02		
bicycle distance * minimize turns	disagree	0,04	0,03		
[ref: preference to minimize turns]	neutral	**0,06***	0,02		
car distance * minimize turns	disagree	-0,00	0,00		
	neutral	**-0,04***	0,00		
	agree	0,04	0,00		
trips by car * active navigation ratio		**0,14****	0,05		
basic education * unfamiliar streets	avoid			**-1,96*****	0,46
	neutral			**-1,01*****	0,35
	attract			**1,32****	0,78
vocational education * unfamiliar streets	avoid			**-1,57*****	0,46
	neutral			**-1,01*****	0,34
	attract			0,02	0,47
university degree * unfamiliar streets	avoid			-0,10	0,58
	neutral			**-0,68***	0,34
	attract			^	
no preference for short cuts * greenopy	avoid			-0,49	1,71
	neutral			**-1,29*****	0,25
	attract			**-0,50***	0,26
neutral to short cuts * greenopy	avoid			-0,76	0,53
	neutral			**-0,44***	0,21
	attract			-0,10	0,24
preference for short cuts * greenopy	avoid			0,25	0,80
	neutral			-0,32	0,24
	attract			^	

N = 1101. ***, **, * Significant at 99%, 98%, 95% confidence level, ^ reference, [blank] not included in model

Socio-demographic. In line with literature, gender has a strong effect. Compared to men, women have more often a self-reported average score, but not necessarily a lower score than average for Orientation Ability (beta-coefficient of −1.36), while the threshold for a low score for Orientation Ability is −1.39 (see Table 3). The odds-ratio implies that compared to men, women have on average 26% chance of having a high self-reported sense of orientation, holding all other variables in the model constant. However, there is also a negative significant interaction effect (-0,08) for men and the average daily distance travelled by bicycle. Therefore, gender differences in bicycling behaviour have been investigated in more detail. A Mann-Whitney U test indicated no significant difference (U = 147.752,5, p = 0,70) between average daily distances travelled by bicycle for women and men. Therefore, it can be concluded that only for men each additional travelled kilometre by bike corresponds to a 0,08-unit decrease in the Logit of the expected value of “high Orientation Ability”, holding all other variables in the model constant. Theoretically, this implies that on average for men who cycle on average 17 km a day, their self-reported Orientation Ability has equal chances for a low, average, or high score, compared to women, holding all other variables in the model constant.

Urban environment. Respondents who are not inclined to make detours to travel along familiar streets have an odds-ratio of 1,5. Hence, they are more likely to self-report a high Orientation Ability compared to respondents that would make detours because they value to travel through familiar streets. The negative effect for respondents who indicated to avoid familiar streets is not significantly different from attraction, which can be explained by the small group size. Overall, these findings suggest that people with lower levels of Orientation Ability compensate for the complexity of the urban wayfinding task by preferring a longer route along familiar streets. Furthermore, it can be hypothesized that high Orientation Ability is more likely to correspond to higher variability in the streets of chosen routes. In literature there are already some indications that also navigational preferences (Hölscher et al., 2011) and salient characteristics of the environment (Li and Klippel, 2016) result in different route patterns. A similar reasoning applies to natural boundaries (rivers and lakes) and perceived accessibility of the bicycle infrastructure in the neighbourhood. Indifference to natural boundaries corresponds to a lower probability to have high self-reported Orientation Ability, and a positive perception towards bicycle accessibility corresponds to a higher probability to have high self-reported Orientation Ability. The latter result implies that higher (perceived) connectivity of the bicycle infrastructure requires more Orientation Ability than average. This is in line with the evolutionary model of Giannopoulos et al. (2014) including the complexity of wayfinding decisions in relation to the complexity of the urban environment, spatial ability, and preferences.

Navigational preferences. Regarding navigational preferences, both bearing line and minimize turns are included as determinants. These preferences are correlated, but including both determinants does not change the direction of the relation with Orientation Ability. The preference in favour of following the bearing line (the perceived direction towards the destination) is associated with a higher probability to self-report a high Orientation Ability, while no preference (or a neutral attitude) to minimize turns is correlated with a higher probability to have high self-reported Orientation Ability. This implies that using these two navigational strategies simultaneously is not beneficial for high self-reported Orientation Ability. Especially respondents that are aware that minimizing turns is not one of their navigational preferences have an increased likelihood to self-report a high Orientation Ability, as their odds-ratio is 2,77.

Interaction effects. Furthermore, there are three significant interaction effects with navigational preferences: average daily distance travelled by car and by bicycle with preference to minimize turns, and the average daily number of trips made by car and active navigation ratio. Compared to respondents with a preference to minimize turns, every additional cycled kilometre of respondents with a neutral attitude to minimize turns corresponds to an increase of 0,06 in the Logit of the expected value of “high self-reported Orientation Ability”, holding all other variables in the model constant. For the average bicycle distance of 3,14 km, the combined beta-coefficient is 0,19. In other words, every additional kilometre cycled a day amplifies the positive effect of a neutral preference to minimize turns.

Contrarily, car distance has a negative coefficient for a neutral preference to minimize turns. The average distance by car reported in the MPN travel diary is 29,5 km, which results in a combined beta-coefficient of −0,72. Thus, it can be concluded that for car travellers, being aware that minimizing turns is not a preference increases the chance to have high self-reported Orientation Ability, while indifference will decrease the chance. Last, there is a negative relation for respondents with low Orientation Ability between the daily number of trips made by car and the active navigation ratio. In other words, for people with lower levels of Orientation Ability the number of trips made by car is higher when on average the respondent is less often in control of the navigation (e.g. as the passenger in a car, or when the daily mobility pattern also includes public transport trips), while for people with high orientation there is positive relation.

2.5.1.2 *Urban wayfinding style II: Knowledge Gathering & Processing Ability*

The second GLM model is used to estimate the latent ability for spatial knowledge. The variables with the highest factor loadings are: regularly choose new routes, exploration attitude to find new routes, and coordinated perception of the environment (NSEW). A total of 6 parameters are included in the GLM (Table 4). The main effects are gender (female or male), age (teenagers, young adults, middle-aged adults, or (young) seniors), navigation preference to follow the bearing line towards the destination (disagree, neutral, or agree), and reported average daily distance travelled by car (0,0–454,0 km). Also, two interaction effects yield significant results; education level and preference to make detours due to unfamiliar streets (detour due to attraction, neutral, or detour due to repulsion). The second interaction is a navigational preference to take short cuts (disagree, neutral, or neutral) and preference to make detours due to the greenopy of the street (detour due to attraction, neutral, or detour due to repulsion).

Socio-demographic. The strongest negative effect on Knowledge Gathering & Processing Ability is found for teenagers (-1,18), with a threshold for an average level of Knowledge Gathering & Processing Ability of −1,27. The odds-ratio implies that, compared to (young) seniors, teenagers have 31% chance of having a high self-reported Knowledge Gathering & Processing Ability, holding all other variables in the model constant. This means that, although significant, age category is not sufficient to distinguish between the 3 levels of Knowledge Gathering & Processing Ability. The main difference in socio-demographic compared to the Orientation Ability model is that the effect of age (odds-ratio of 0,31) is stronger than the effect of gender (odds-ratio of 0,51).

Navigational preferences. Regarding navigational preferences, the preference for following the bearing line (or direction towards the destination) corresponds to a lower probability to have high self-reported Knowledge Gathering & Processing Ability, while a neutral attitude to minimize turns corresponds to a higher probability to have high self-reported Knowledge Gathering & Processing Ability. Note that this determinant has the opposite effect on Orientation Ability. Hence, people reporting a lower Knowledge Gathering & Processing Ability and a higher Orientation Ability are more likely to correspond to a navigational preference to follow the bearing line. This shows that some determinants have an ambiguous effect on both wayfinding styles; also car distance, included as interaction effect for Orientation Ability, has a positive relation with Knowledge Gathering & Processing Ability, and a negative relation with Orientation Ability. Each additional travelled kilometre by car corresponds to a 0,01-unit increase in the Logit of the expected value of "high Knowledge Gathering & Processing Ability" versus lower conditions, holding all other variables in the model constant.

Interaction effects. Furthermore, there are two interaction effects with the attraction to urban elements while using the bicycle for a personalized trip purpose. Exploration of unfamiliar routes has an interaction effect with the highest completed education level. Significant differences with people with high education and attraction to unfamiliar routes are negative for people with basic or vocational education and repulsion or neutral attitude towards unfamiliar streets. For people with a university degree, there is only a smaller significant negative difference for neutral attitude towards unfamiliar streets. Being attracted to unfamiliar streets is not significantly different between people with a university degree and vocational education, but there is a higher chance for people with a low vocational education (5% of the respondents) to have a significantly higher self-reported Knowledge Gathering & Processing Ability. One reason could be related to differences in mental, verbal and memory abilities also found in (Shelton et al., 2013). Another possibility is that children younger than 21 have no chance of having a completed vocational education or university degree and therefore are compensated if they state to be attracted by unfamiliar streets. A third reason could be to latent different travel patterns between different levels of education. However, the current data and model do not provide strong evidence for these explanations.

The second interaction effect is the navigational preference to take short-cuts and the greenopy (street with trees). What can be observed is that with a preference to take short-cuts, effect of the attitude towards greenopy is not significantly different on their self-report Knowledge Gathering & Processing Ability. Thus, taking short-cuts in very urban (many buildings, little green) areas and areas where trees have a prominent role requires similar Knowledge Gathering & Processing Abilities. However, as merely 1% of the respondents indicated to avoid streets with trees, the results do indicate that avoiding greenopy requires a little more Knowledge Gathering & Processing Ability. For respondents who do not prefer to take short-cuts, cycling a longer distance due to attractive greenery corresponds with lower Knowledge Gathering & Processing Ability. This could be interpreted as a detour through green streets is easier to memorize for people that do not wish to make short-cuts along the route. Secondly, with neutral or no preference to take short-cuts and a neutral attitude towards the greenopy, the chance to also have a lower level of Knowledge Gathering & Processing Ability is higher. 40% of the respondents indicate to be attracted to make a detour along green passages to the activity they most frequently visit by bicycle. Similar are natural boundaries caused by rivers and lakes beneficial and they do not require more Orientation Abilities. Also, in this chapter the perceived accessibility of the bicycle infrastructure showed a significant relation with Orientation Ability, which is in line with existing literature where street connectivity is a significant determinant. These results could be important insights for the design of active and healthy cities; each additional minute travelled by foot or bike can beneficial for somebody's health. However, more research is needed to investigate how urban design affects the number of bicycle trips, bicycling time and distance.

2.5.2 Model evaluation

This section evaluates the two estimated GLM models. Contingency tables (Table 5 A, B) are used to calculate the prevalence (overall accuracy). It is defined as the number of all correct predictions divided by the total number of respondents from the contingency tables, with an evaluation of 1 (0) as the best (worst) possible.

Table 5 A-C. Contingency tables of wayfinding styles.

A. Orientation Ability

Predicted	Actual			Total
	-1	0	1	
-1	**168**	112	41	320
0	124	**149**	127	396
1	41	136	**203**	385
Total	333	397	371	***1101***

B. Knowledge Gathering & Processing Ability

Predicted	Actual			Total
	-1	0	1	
-1	**137**	110	24	237
0	168	**243**	174	641
1	30	83	**132**	223
Total	335	436	330	***1101***

C. Score

Weights			
5	1	0	**6**
1	4	1	**6**
0	1	5	**6**
6	**6**	**6**	

On average both Orientation Ability and Knowledge Gathering & Processing Ability yield a prevalence of 0,47. These results are acceptable as they are a 42% improvement compared to a random accuracy of 0,33. However, these evaluations can be too optimistic when the accuracy of the prediction is unequally distributed. An estimate that is only one degree off (predicted "average ability" instead of "low" or "high") is better than an estimate predicting a "high ability", while it should have been a "low ability" (second degree). Therefore the models are also evaluated with a weighted scoring shown in Table 5C. This particular combination of weights yields a maximum score of 5108 points for 1.101 respondents, with an evaluation of 1. The "penalty" is higher when the prediction is 2 conditions of (0 points) compared to 1 (1 point), while there is no differentiation between conditions (each row and column has equal points; 6). The model describing Orientation Ability yields a score of 0.57, while Knowledge Gathering & Processing Ability performs slightly less with a score of 0.56.

Each contingency table can be used to derive three confusion matrices for a more detailed evaluation (Table 6). For each of the three ability conditions (lower than average (−1), average (0), higher than average (1)) a confusion matrix is computed that reports the number of false positives, false negatives, true positives, and true negatives. The sensitivity is calculated as the number of correct positive predictions divided by the total number of actual positives. Colloquially, given a specific ability condition (i.e. low, average, or high) how often is the prediction correct. From Table 6 it can be observed that given a condition both wayfinding models yield predictions that are for 38% to 56% of the cases correct. "Average Orientation Ability" and "high Knowledge Gathering & Processing Ability" have the lowest performance, while "average Knowledge Gathering & Processing Ability" and "high Orientation Ability" have the highest. Regarding specificity, the table indicates that only "average Knowledge Gathering & Processing Ability" scores low. In other words, the models are quite suitable to identify a respondent that has not low Orientation Ability as someone who has either neutral or high Orientation Ability

Table 6. Evaluation measures based on derived confusion matrices

		Orientation Ability			**Knowledge Gathering & Processing Ability**		
		Condition i			*Condition i*		
Description	**Formula**	-1	0	1	-1	0	1
Prevalence	$TP_i / (TP_i + FN_i + FP_i + TN_i)$	0,30	0,36	0,34	0,30	0,40	0,30
Sensitivity (TPR)	$TP_i / (TP_i + FN_i)$	0,50	0,38	0,55	0,41	0,56	0,40
Specificity (TNR)	$TN_i/(TN_i + FP_i)$	0,80	0,64	0,76	0,83	0,49	0,85
Precision (PREC)	$TP_i / (TP_i + FP_i)$	0,52	0,37	0,53	0,51	0,42	0,54
Informedness	$TPR_i + TNR_i - 1$	0,31	0,02	0,30	0,23	0,04	0,25
Description	**Formula**	Weighted average			Weighted average		
Overall accuracy	$(TP_1 + TP_2 + TP_3) / (P + N)$	0,47			0,47		
Weighted score	$((TP_1 + TP_3) * 5) + (TP_2 * 4) + FN^1 + FP^1 / ((AA_1 + AA_3) * 5) + (AA_2 * 4)$	0,57			0,56		
Overall informedness	$TPR_{w.mean} + TNR_{w.mean} - 1$	0,21			0,20		

P: all Positive, N: all Negative, TP: True Positive, FN: False Negative, FP: False Positive, TN: True Negative, FN^1: first degree False Negative, FP^1: first degree False Positive, AA: Actual Ability.

Moreover, the model precision is calculated as the number of correct positive predictions divided by the total number of positive predictions. This metric is useful if the model is generalized. It provides an indication of how many of the predicted ability conditions are actually correct. For these models, all low and high abilities can be estimated with more than 50% precision. However, the models have a tendency to assign individuals that are "low" or "high" to the "average" ability level category.

The ratio between TPR and TNR also provides insight into any bias to specific ability conditions and how conservative the model is. The TPR and TNR values for orientation abilities are very similar, but there is a bias for average Knowledge Gathering & Processing Ability, while low and high knowledge gathering & processing abilities are too conservative. Finally, Informedness describes the extent of any form of guessing of an informed decision. A value of 0 (both "average" conditions) depicts the highest possible probability that the model outcome is more a guess than an informed decision. From the results in Tables 4 A-B it can be concluded that the quality of the model results is not equally distributed. It can be concluded that "low" and "high" wayfinding abilities are better modelled compared to the respective "average" wayfinding ability.

2.6 Synthesis on wayfinding styles

This chapter aimed to investigate differences in urban wayfinding behaviour and relations with individual navigational preferences in the larger (metropolitan) urban environment where daily travel behaviour takes place in The Netherlands (See Figure 8). This section elaborates how the findings of holistic GLMs on Orientation Ability and Knowledge Gathering & Processing Ability contribute to (i) the understanding of urban travel and mobility behaviour, (ii) provision of comprehensible travel information, (iii) design of legible cities, (iv) identify potential navigation problems, and (v) limitations of this chapter.

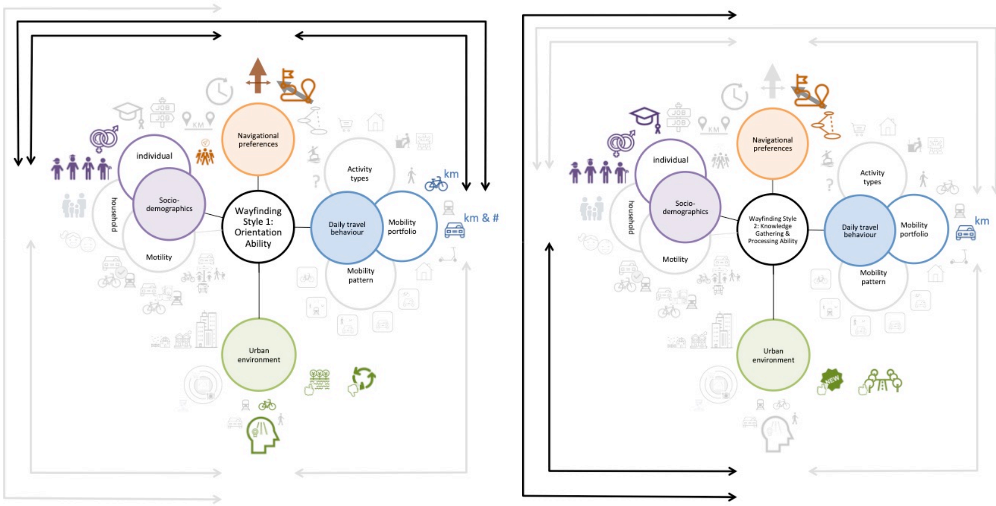


Figure 8. Significant determinants of two urban wayfinding styles in The Netherlands. Dark arrows: significant interaction effects. Urban elements to avoid or attract are rivers and lakes, (un)familiar streets, and greenopy.

2.6.1 Relation between travel behaviour and urban wayfinding styles

From the literature background it was hypothesized that the total average travel distance (by car and foot) have a negative relation with the wayfinding score. The results in this chapter also show a negative relation for distance travelled by car, and for the first time, also distance bicycled by men with Orientation Ability. However, this chapter also shows that the total average distance travelled by car and the interaction effect between average number of car trips and active navigation ratio have positive relations with Knowledge Gathering & Processing Ability. Although the majority of the research found in literature investigates pedestrian wayfinding, the distance travelled by foot and public transport are not significant in this chapter. Furthermore, a latent cluster analysis has been performed using the average number of trips per travel mode (Ton et al., 2019). Noteworthy, although significance in a categorical model only including travel behaviour, in both GLMs the identified mobility patterns clusters did not yield any significant relation in combination with other determinants.

2.6.2 *Travel information and route choice*

Based on findings in literature, it can be expected that respondents with a better sense of orientation choose routes with shorter travel distance and time, but not necessarily higher travel speed. This requires flexible navigational preferences as the structure and layout of each urban environment demands different abilities. However, both GLMs did not include navigational preferences to minimize travel distance or time. Regarding the provision of comprehensible travel information, this indicates that wayfinding styles are more related to number of turns, bearing line and short-cuts than travel distance or travel time. In the future, a similar study including travel data at route level could be used to investigate differences in route choice behaviour and variability.

2.6.3 *Legible urban wayfinding*

Figure 7 demonstrates that approximately 30% of the respondents use the bicycle for urban short trips up to 10 km, while only 15% use the car. This travel behaviour is typical for the Netherlands, where many people consider the bicycle as the main transport mode, especially within cities. Also, the bicycle is an important transport mode to achieve the climate goals stated in the Paris Agreement. Notwithstanding, little is known about what makes it easy to navigate a city by bicycle and how the urban environment affects bicycle behaviour. To this end, this chapter identifies several factors concerning the design of legible cities for cycling behaviour.

Based on the models it seems that a combination of high Orientation Ability and Knowledge Gathering & Processing Ability will correspond to higher variability in the streets of chosen routes. With higher (perceived) connectivity of the bicycle infrastructure more Orientation Ability is required than average. This implies that people with lower levels of Orientation Ability will compensate for the complexity of the urban wayfinding task by preferring a longer route along familiar streets. Thus, even if high connectivity exists, but all people have low orientation abilities, still not much route variation will occur and it will become more difficult to mitigate congestion and distribute large cyclists flows more evenly. Insights related to navigational preferences and urban environment on Knowledge Gathering & Processing Ability can be interpreted as for people that do not wish to make short-cuts, for example due to absent time pressure, it is easier to memorize a detour through a green passage. Last, although urban density has been identified as important characteristic for salience and legibility of an environment, its role as a determinant remains unknown, as neither model indicated significance.

2.6.4 *Interaction between urban wayfinding styles*

Both wayfinding styles can be used complementary as different processes influence them. However, two determinants (navigational preference to follow the bearing line and average daily distance travelled by car) have an ambiguous effect on both wayfinding styles. This could indicate a trade-off, because gathering and processing more spatial knowledge will ultimately require more orientation ability in order to process the knowledge into useable wayfinding styles. The navigational preference to follow the bearing line is not beneficial when there is a low amount of spatial knowledge, as this does not encourage the acquisition of more spatial knowledge. If a satisfactory amount of spatial knowledge has been acquired using the bearing line as a navigational preference is useful to reduce the workload.

2.6.5 *Limitations of urban wayfinding styles*

One of the limitations of this chapter is the assumption that urban wayfinding styles are static personality traits. To investigate if this assumption is valid, either a study should target visitors unfamiliar with a city, or this questionnaire should become part of the longitudinal data collection efforts of the MPN. In the latter case Generalized Estimating Equations (GEEs) can be used to deal with correlated observations, such as clustered data of subjects or classes (Hardin and Hilbe, 2012, Ballinger, 2004). A second recommendation for future work is the development a route choice model including the taste heterogeneity based on wayfinding styles to describe variability is chosen street segments.

The second limitation relates to the subjective nature of factor analysis and self-reporting behaviour of respondents (Fabrigar et al., 1999, Willis et al., 2009). There are indications of socio-cultural differences in reporting behaviour. For example, there might be some variation in how people assess their ability. So far it is unknown to what extent does this depends on the perceived ability of a partner, parents, and/or friends. Additionally, mistakes can be made while completing the three-day travel diaries. Therefore it is recommended to compare the accuracy of travel diaries with activity data and complementary travel data using GPS or mobile phones.

Furthermore, it should be stressed that socio-demographic differences in wayfinding styles in this chapter should be interpreted in terms of variations in development of beliefs and behaviour rather than overall ability or intelligence. Significant gender differences in favour of men are found for both Orientation Ability and Knowledge Gathering and Processing Ability. This is to some extent different from findings in most studies, where the majority of the findings lean towards the hypothesis that men are better at orientation and navigational tasks, while women have enhanced knowledge gathering, memory and processing ability. The difference can be partly ascribed to different questionnaires and experimental set-ups to measure knowledge gathering and processing abilities, as well as measured at different levels of spatial scale (e.g. toy model, indoor, route level, small VR environment, real city, realm of daily travel patterns).

In addition, there are always limitations to the length of a survey. The PAW-AM data collection is designed to gather many insights regarding pedestrian and cyclist mobility behaviour, from attitude towards mode choice, social norms, to wayfinding behaviour. Consequently, to reduce respondents' fatigue a number of questions related to avoidance and attraction of urban element was limited to bicycle trips. As this chapter is the first of its kind to investigate these types of bicycle landmarks, it is recommended to extend this approach to other travel modes to capture the complete picture of urban legibility.

Finally, with extensive models this chapter shows that only a limited number of determinants have a combined effect on each wayfinding style. Although more determinants, such as the mobility cluster patterns, show significant relations if included in solitude, it is believed that these simple model results are too optimistic. Moreover, many variables have been included in the surveys and have been tested for in the GLMs, many variables still need to be investigated. Both reasons probably contribute to the relative low Accuracy and Informedness of both models.

2.7 Conclusion

This is one of the first studies to investigate differences between urban wayfinding styles in relation to travel behaviour and navigation preferences in The Netherlands. Dutch travel behaviour is rather particular with relative short travel distances, a substantial amount of intercity commute, a long history of high bicycle shares, many rivers and canals, and nearly no inclination. Therefore, moderate differences with existing studies are expected and a Confirmatory Factor Analysis can be used for the generalization of the content of wayfinding styles in other contexts and estimate to what extent similar determinants have an influence.

The main contribution of this chapter is the theoretical insight of how urban wayfinding behaviour relates to daily travel patterns. Moreover, possibilities and relevance for route choice behaviour, identify potential navigation problems, design more legible cities, and provision of comprehensible travel information are discussed.

Two holistic Generalized Linear Models (GLMs) describe urban wayfinding styles based on two dependent factor components "Orientation Ability" and "Knowledge Gathering & Processing

Ability”. The results are acceptable as they are a 42% improvement compared to a random accuracy of 0,33. However, the quality of the model results is not equally distributed; “low” and “high” wayfinding abilities are better modelled compared to the respective “average” wayfinding ability. The following determinants are significant: gender, age, education level, perceived bicycle accessibility of the neighbourhood, attraction to familiar and unfamiliar streets, and greenopy of the streets, navigational preferences to minimize turns, follow the bearing line, and take short-cuts, ratio of active navigation, average daily distance travelled by car and bicycle, and average daily number of trips made by car. Gender and age have similar effect signs on both OA and KA, while the navigational preference to follow the bearing line and average daily distance travelled by car have disassociated effects. The remaining determinants are only significant in either OA or KA, providing evidence that predominantly different processes describe each wayfinding style.

Acknowledgements

This research was supported by the Allegro project (no. 669792), which is financed by the European Research 8 Council and Amsterdam Institute for Advanced Metropolitan Solutions. The data was made available by the Netherlands Mobility Panel administered by KiM Netherlands Institute for Transport Policy Analysis.

Use this map to explore, indulge, relax and discover the sites of your own City!

KEY

Selfie spot
Street art

ADELAIDE FREE BIKES

Hire an Adelaide FREE Bike from any of the 21 locations across the City and North Adelaide including Adelaide City Council, Adelaide Zoo, North Adelaide Golf Course and more. All you need is your current driver's license, proof of age card or passport.

bikesa.asn/au/AdelaideFreeBikes

FREE CITY CONNECTOR BUS

The free City Connector bus service runs on two loops – an inner city loop and an extended loop around North Adelaide providing a link to popular City attractions and shopping, dining and services destinations.

adelaidemetro.com.au/freecitybus

TOURS

Experience Adelaide on a guided tour, visit our website for more information on many of the City's fantastic tour providers.

PARKING

Driving to the City? Park at one of UPark's 9 convenient locations some of which offer overnight parking and online bookings.

upark.com.au

Download your itinerary at
cityofadelaide.com.au/beatourist

Share your journey
#beatouristadl #myadelaide

Use this map to explore, indulge, relax and discover the sites of your own City!

KEY

Selfie spot

Street art

ADELAIDE FREE BIKES

Hire an Adelaide FREE Bike from any of the 21 locations across the City and North Adelaide including Adelaide City Council, Adelaide Zoo, North Adelaide Golf Course and more. All you need is your current driver's license, proof of age card or passport.

bikesa.asn.au/AdelaideFreeBikes

FREE CITY CONNECTOR BUS

The free City Connector bus service runs on two loops – an inner city loop and an extended loop around North Adelaide providing a link to popular City attractions and shopping, dining and services destinations.

adelaidemetro.com.au/freecitybus

TOURS

Experience Adelaide on a guided tour, visit our website for more information on many of the City's fantastic tour providers.

PARKING

Driving to the City? Park at one of UPark's 9 convenient locations some of which offer overnight parking and online bookings.

upark.com.au

Download your itinerary at
cityofadelaide.com.au/beatourist

Share your journey
#beatouristadl #myadelaide

3 Spatial Analytics for Identification of Salient Areas

Spatial urban route knowledge consists of the internalized representation of a sequence of actions to be performed at certain locations, cued by wayfinding landmarks. Determining the location of distinctive landmarks is thus important in research on route choice, urban cognition, and travel information. Currently, most approaches to identify landmarks require vast data collection efforts. To overcome these demands, this chapter proposes a spatial analytic method able to handle open-source datasets to identify urban wayfinding landmarks as salient urban areas.

1. Introduce aggregate urban landmarks (salient urban areas) as noticeable areas with distinct characteristics from their (local and/or global) surroundings.
2. Developped a spatial analytic methodology to identify urban salient areas based on open-source GIS data (BAG and GBKA).
3. Relevant determinants are building volume, surface, height, building year, and the number of buildings in a 100 m^2 grid-cell.
4. Gini coefficient unravels differences in spatial distribution of clustering and dispersion of urban salient areas.

This is an edited version of the following article:
Zomer, Duives, Cats, and Hoogendoorn (2017). *Spatial Analytics for Identification of Salient Areas*. Paper presented at Conference Urban Planning and Urban Management (CUPUM*)* 2017, Adelaide.

3.1 Introduction

City users, to some extent, rely on memorized urban route knowledge to decide how to move from one place to the next. To this end, spatial urban route knowledge can be viewed as remembered sequences of landmarks, that, combined with directional actions support users to navigate across town. Following Lynch (1960) and Appleyard (1970), landmarks are defined as salient geographic objects, points, or polygons of buildings that structure the internal representation of a city (Richter and Winter 2014).

Over the last two decades, different approaches to identify and integrate landmarks have been developed, as can be noticed, e.g. in route descriptions. As such approaches require large-scale, detailed, diverse datasets, and correspondingly demanding data collection methods (Richter and Winter 2014), today, knowledge on the effects of urban landmark distribution on wayfinding behavior remains limited.

This chapter aims to contribute to methodology with an approach to handle open-source data. To do so, the concept of aggregate urban landmarks, coined as salient urban areas, is introduced. Salient urban areas possess noticeable characteristics that make them distinct from their surroundings. From a theoretical perspective, a landmark is salient (distinct) in relation to its immediate surrounding or context at large. Salient urban areas are considered unique, either because of dissimilarities to their (local) area, and/or else, because of characteristics considered similar in comparison to other (global) areas. Presumably, the more distinctive a landmark or area, the easier it will be to memorize and incorporate this saliency into the spatial route knowledge to be drawn upon in future. Therefore, salient urban areas are hypothesized to be important to structure spatial knowledge in longterm memory (Couclelis et al. 1987; Sadalla et al. 1980; Montello 1997).

Any method to identify salient landmarks has to be applicable in largescale environments presenting unequal distributed data. Using open-source data on Amsterdam's urban structure, this chapter examines whether a spatial analysis approach is useful to identify salient urban areas. First, determinants in urban environments will be defined. As previous studies focused on identifying and integrating landmarks as salient buildings, metrics for salient urban areas can be inferred. Next, to allow for a systematic analysis, a cellular grid (100 square meters per grid-cell) is projected covering the case study area in Amsterdam. Last, grid-cells' determinants are spatially analyzed to identify the characteristics of local hotspots and global clusters of salient urban areas.

Section 3.2, synthesizing prior studies, offers insights into the identification of urban landmarks and methods to conduct research on landmark identification in relation to wayfinding behavior. Section 3.3 elaborates on the research approach and methodology. Results, presented in section 3.4, are categorized into three subsections, starting with a descriptive analysis of the determinants used for the Amsterdam case study. Next, both the findings on the identification of salient urban areas and analyses on the (spatial) distribution of identified landmarks will be put forward. Section 3.5 summarizes the conclusions and provides recommendations for further research.

3.2 Defining urban landmarks

In this chapter, the literature review on the influence of urban structures on city users' wayfinding behavior focuses on landmark identification and urban typologies. To this end, section 3.2.1 presents first insights derived from cognitive sciences regarding landmark identification, and, next, in section 3.2.2 urban morphology techniques to distinguish urban typologies are put forward.

3.2.1 *Identifying landmarks*

The concept of landmarks originates from Lynch's research (1960) in which five elements according to which cities are perceived, comprise paths, nodes, landmarks, edges, and areas. Appleyard (1970) combines landmarks, being both objects in space and internal representations, with the notion of salience and hypothesizes the more unique a building is, the more likely it will be incorporated into survey knowledge.

Based on memorized buildings in one's "home town", Appleyard identifies significant determinants, both for local (neighborhood) and global comparison (across city areas) based on memorized buildings in a "home town". Using the correlation between property and frequency of recall, the author distinguishes three properties: form (contour, building volume, visual attributes of the façade), semantics (intensity and uniqueness of use), and structural (location and structure of environment). Resulting from Hillier's and Hanson's space syntax theory (1984), a fourth property, visibility (frequency of being in-sight and proximity to a vantage point), has been added (Morello and Ratti 2009). Working on the isovists idea regarding visible sights, Morello and Ratti argue urban environments will be legible due to their location-based visibility. Richter and Winter (2014) hold a building's total salience to become stronger as its distinctiveness on more categories increases.

It appears, whereas in urban planning, landmarks appear firmly grounded concepts, their appliance to large-scale environments is cumbersome, particularly, when buildings are unequally distributed. Based on Lynch, regarding their identification, generally, landmarks are analyzed as georeferenced points or buildings. Although, resulting from social data, using pictures, new approaches to identify landmarks from (geo-referenced) user-generated data are being developed (Duckam et al 2010; Richter 2007), an aggregated, cellular approach is still lacking.

3.2.2 *Landmarks determinants based on urban morphologies*

Landmark identification frameworks appear intricate to apply to spatial experiences (Stevens 2004). This may be one reason why systematic research on how people learn and comprehend novel urban environments – i.e., how people organize, group, differentiate and catalogue their perceptions while moving across town – remains limited. Urban morphology aims to understand spatial structures and patterns, e.g., physical layouts of urban environments, but its underpinning methods have not been applied to identify salient landmarks in urban environments.

Levels of analysis in urban morphology range from regional to continuous points. Main objects of interest are building blocks, followed by neighborhoods. As can be noted in Table 7, in urban morphology, determinants may be quantified along different scales, and, moreover, depending on particular research goals, decisions as to what determinants to include, may vary. The earliest methods distinguish typologies (urban atmospheres) based on conceptual differences (Lynch 1960; Conzen 1960; Duany 2002; ABF Research 2003). From 2000, different methods have been used, such as plotting against two axes (Marshall 2005; Berghauser-Pont and Haupt 2010), or by applying a characteristic, to some extent (Morello and Ratti 2009; van Nes et al 2012; Oliveira and Medeiros 2016).

We conclude that although various methods and techniques have been developed to identify landmarks in relation to wayfinding behavior, little is known on how the distribution of landmarks in large-scale urban environments actually effect wayfinding behavior. From literature, shape turns out to be a consistent indicator in both landmarks and urban morphology, and, therefore, urban grid-cell landmarks will be identified using aspects of shape.

Table 7. Landmark determinants in urban morphology.

Scale	Determinants
Regional	City size Density (FSI, GSI) Proximity of services Land use mixture Building period
Plot or neighborhood	Town plan Land use pattern Composition of network hierarchy and directionality Configuration of intersection and connectivity of network Betweenness centrality Building density (FSI, GSI)
Street	Average (pedestrian) flow
Building block	Density and volume of the built environment (spacematrix) Land use mixture Building form pattern
Grid	Spatial integration of axial lines Building densities (FSI, GSI) Land use mixture
Continuous	Accessibility of network Ground Space Index (GSI) Building year Mixed building usage

3.3 Research approach and methodology

This section, first, introduces the research approach, followed by an explanation of data processing procedures and spatial analyses using ArcGIS to analyze salient urban areas. Last, the case study area and cleaning processes on open-source data will be discussed.

3.3.1 Research approach

Spatial route knowledge on a city can be conceived of as the cognitive level of route choices, consisting of memorized (orders of) landmarks. It is hypothesized for landmarks characterized by more noticeable local or global (dis)similarities to be easier to memorize, and, thus, to be more probable to become part of the cognitive level of route choices. Following Lynch, in order to be distinct from its nearby surroundings, a landmark is to be strongly dissimilar from local buildings. Likewise, clusters (neighborhoods) can be considered distinct when there is a strong global similarity in terms of continuity and delineation of space. As form turns out a consistent characteristic, both regarding landmarks and urban morphology, urban grid-cells are identified by contour (Ground Space Index – GSI), volume (Floor Space Index – FSI and number of floors – L), and visual attributes of the façade (building year).

3.3.2 *Data processing procedure and analysis*

As concluded in section 3.2, due to approaches requiring vast data collection efforts, knowledge on the effects of the distribution of landmarks in large-scale urban environments on wayfinding behavior remains limited. To try and fill this gap to some extent, we propose a spatial analysis method to translate detailed disaggregate data of large-scale urban environments into meaningful and computationally efficient aggregate data. Below, five steps comprising the spatial analysis method are introduced. In 3.3.3.1, these methodical steps will be applied to the case study.

Step 1. Create map layers from data.
  a. Create grid-cells using a fishnet that superimposes the area of interest.
  b. Assign available data to grid-cells.

Step 2. Iterative grouping analysis to identify how the determinants relate to different urban morphologies.

Step 3. Cluster and outlier analysis based on Anselin Local Moran's I using the determinants of interest as input fields (Anselin 1995).
  3.1 An inverse distance squared is used because nearby neighboring grid-cells have a much larger influence than grid-cells further away.

Step 4. Unite map layers of cluster and outlier results of relevant determinants.

Step 5. Create maps of urban salient areas.
  5.1 Local urban salient areas: Cumulative summation of all low negative z-scores indicate statistically significant spatial outliers of a high value surrounded by low values (HL) and a low value surrounded by high values (LH).
  5.2 Global urban salient areas: Cumulative summation of all high positive z-score that indicate statistically significant clusters of high values (HH) and low values (LL).

3.3.2.1 *Discussion of the elements of the spatial analytic method*

Upon the creation of grid-cells, it has to be ensured such cells are of large enough size to contain at least one feature, and small enough to allow for variety within urban plots. Also, the size of grid-cells should be suitable for further analysis. Furthermore, dependent on available data, the specific combination of spatial joints, intersections, dissolves and unions to be used to transform the data to grid-cells will have to be decided.

Grouping analysis is used as an exploratory analysis to reveal underlying structures of the determinants of interest to be used to identify clusters of distinct urban areas with similar physical characteristics (Jain 2009).

Cluster and outlier analysis is applied in many domains, such as economics and geography to identify concentrations of values and outliers that explain (behavioral) patterns (Anselin 1995). This analysis is often preferred over hotspot analysis based on the Getis-Ord Gi*, as it also identifies statistically significant spatial outliers, which are expected to be the most important aggregate urban landmarks.

3.3.3 *Amsterdam as a case study*

Next, the case study area is presented as well as the operational choices needed to apply the spatial analytic method to the case study.

Founded in the 13th century, Amsterdam is situated along the river "het IJ" and the Amstel delta. Following several expansion periods, Amsterdam's residential area covers over 165 km^2. The open-source GIS data provided by the City of Amsterdam to analyze the urban structure of Amsterdam can be downloaded at http://maps.amsterdam.nl/open_geodata. This dataset consists of 471.580 BAG (key registers of addresses and buildings) address points that include attributes concerning building year and usage surface, and 17.791 polygon shapes of

GBKA building typology (large scale topography), by which the surface area can be calculated to represent the Ground Space Index (GSI). Due to missing and incorrect geo-coded data, 23.532 (5%) points of the BAG addresses have been excluded because they did not include a building year, and 17.021 (4%) have been excluded because they did not include

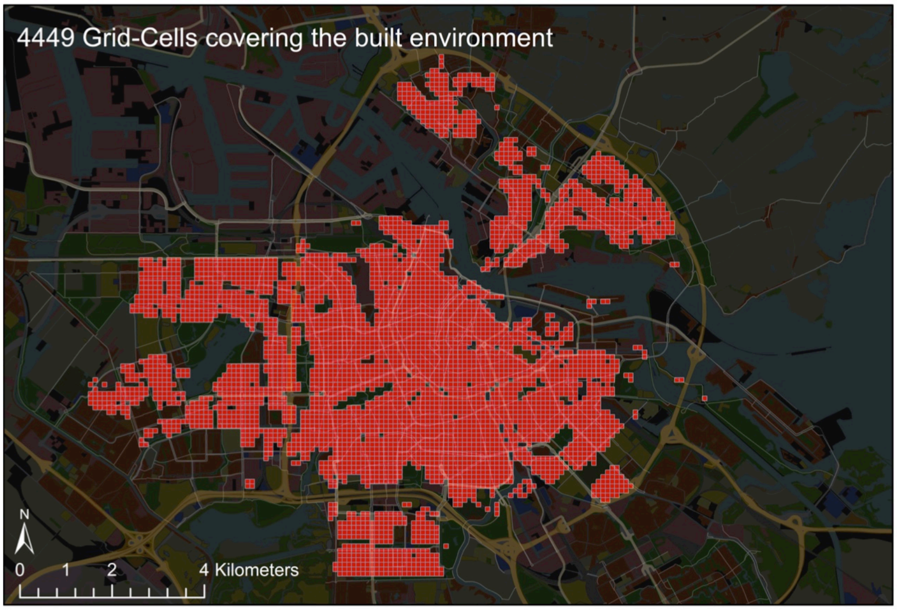


Figure 9. Case study Amsterdam. In red: all $100m^2$ grid-cells with one or more built environment features.

a usage surface. Regarding the polygons of buildings from GBKA, 13.437 (75%) polygons containing a BAG address point together with a building year and/or usage surface have been included.

3.3.3.1 Applying five steps of the analytic method to Amsterdam

Step 1. To ensure that most cells are large enough to contain at least one feature, and, simultaneously, small enough to allow for diversity within urban plots, the area of the grid-cells is set at 100 square meters. After data cleaning and overlaying open-source point data on BAG addresses with the polygon shapes of corresponding buildings, Floor Space Index (FSI), Ground Space Index (GSI), and number of floors (L) are calculated. As a direct spatial joint of many-to-one does not exist in ArcGIS, several steps (intersections, spatial joints and dissolves) are necessary to link the processed open-source data sets to the grid-cells. The final step in data processing combines all processed layers to one, which can be used for spatial analyses. This resulted in “aggregated” data for 4.449 grid-cells, each containing at least one building, and a maximum of 73 buildings, see Figure 9.

Step 2. In deciding which determinant to include, a first selection is made based on the semantic meaning in relation to the concept of aggregate urban landmarks. Secondly, meaningful determinants should have at least one determinant where similar values are dispersed, and one determinant where similar values are clustered (referring to local and global salient urban areas). Table 8 gives an overview of these intermediate steps as well as 12 determinants that are hypothesized to describe the spatial pattern of salient urban areas. N/A indicates the determinant is either not applicable or there is no assignment; Low or High indicate whether low or high values are spatially dispersed (local) or clustered (global). The blank fields indicate a random distribution of the values of meaningful determinants.

Table 8. Urban landmark grid-cells identification metrics.

	Determinants	Count	Sum	Min	Max	Range	Std
	Building Year	N/A	N/A		N/A	High	High
	Building Volume (FSI)	N/A		High			
Local	Contour Surface (GSI)	N/A		High			High
	Levels	N/A	N/A	N/A	High		
	Number of buildings per grid-cell		N/A	N/A	N/A	N/A	N/A
	Building Year	N/A	N/A	Low	N/A		Low
	Building Volume (FSI)	N/A	Both				Low
Global	Contour Surface (GSI)	N/A	Both				Low
	Levels	N/A	N/A				Low
	Number of buildings per grid-cell	Both	N/A	N/A	N/A	N/A	N/A

Grouping analysis aims to explore spatial patterns and identify the reliability of the 12 determinants hypothesized to describe these patterns. When performing this grouping analysis, each determinant has a Rho^2, describing the extent of discrimination amongst determinants. Because there is no ground truth about either the determinants or the identification of groups, a suitable determinant is defined as a determinant with a low range of Rho^2 for different number of groups. Furthermore, in grouping analyses no spatial constraint is used; features are partitioned using a k-means algorithm to minimize differences amongst features in a group, over all groups. Multiple iterations have been performed to identify suitable combinations of determinants to overcome the limitations of the greedy heuristic.

Step 3. Based on Anselin Local Moran's I statistic (Anselin 1995), cluster and outlier analysis identify statistically significant hot and cold spots and spatial outliers. Incremental spatial autocorrelation analyses provide insight into the maximum spatial autocorrelation. However, for many determinants the distance band turned out too high to ensure that no feature exceeds 1.000 neighbors, which results with memory errors. Therefore, the fixed distance band was set at 700 meters. All grid-cells within the distance band are weighted equally.

Step 4-5. Final maps of urban salient areas can be created when the results of the cluster and outlier analysis are combined with "union". The total level of salience of a grid-cell is the cumulative score of significant values. Significant values of low negative z-scores of suitable determinants are summed to represent local salient urban areas. Significant values of high positive z-scores of suitable determinants are summed to represent global salient urban areas.

3.4 Results

Section 3.4, first, discusses descriptive statistics, followed by the identification of local and global salient urban areas. Finally, in the last part, a possible application of the spatial analysis approach is discussed, aiming to investigate the spatial distribution of salient urban areas with the Gini coefficient.

3.4.1 *Descriptive statistics on the case study Amsterdam*

Descriptive statistics regarding the case study Amsterdam are shown in Table 9. The oldest buildings in the dataset date stem from 1300, and the average age of buildings within a grid-cell is 35, with a maximum of 709 years. On average the built volume of a grid-cell is 5.145 m^3, with a maximum of 110.288 m^3. If the surface of the buildings would be 10.000 this would correspond to 10 floors. On average the surface of buildings cover almost 20% of grid-cells. Within a grid-cell, the average smallest surface equals 343 m^2, whereas the maximum equals 10.000 m^2. The highest building level within a grid-cell reaches almost 23 floors, while the average building level is below 3 floors. The average value is lower than expected for an urban area like Amsterdam, and probably, results from an incomplete dataset. The average number of buildings within a grid-cell is just over 6, with a maximum of 73.

Table 9. Descriptive statistics on determinants for case study Amsterdam. In bold: determinants of interest.

	Determinant	Mean	Std. Dev.	Min	Max
Building year	**Oldest** (min)	1912	78,91	1300	2016
	Newest (max)	1947	33,50	1600	2016
	Range	35,86	81,56	0	709,00
	Std. Dev.	13,61	29,03	0	251,25
FSI	**Average**	5.145,64	4.523,64	0	110.288
	Smallest (min)	879,27	2.675,77	0	110.288
	Largest (max)	2.521,82	3.120,71	0	110.288
	Std. Dev.	737,02	1.229,74	0	5.176,39
GSI	**Average**	1.830,73	1.242,90	0	10.000
	Smallest (min)	343,73	722,34	0	10.000
	Largest (max)	911,78	769,25	0	10.000
	Std. Dev.	256,48	311,93	0	3.971,81
Building level	Average	2,76	1,38	0	22,86
	Lowest (min)	2,04	1,45	0	22,86
	Highest (max)	3,60	2,04	0	22,86
	Std. Dev.	0,58	0,74	0	16,16
Number of buildings	**Count**	6,28	7,97	1	73

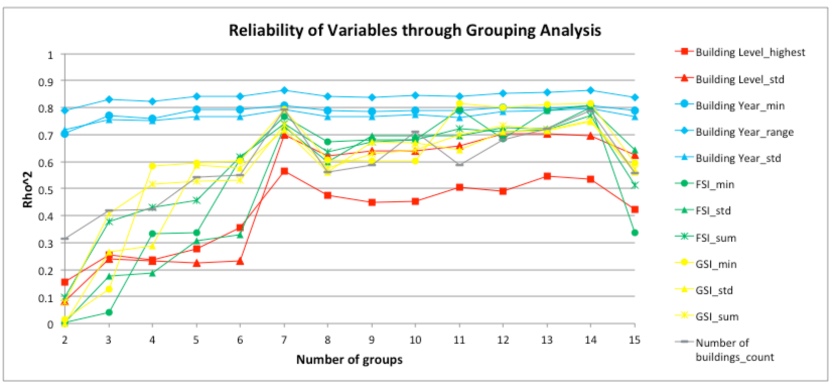


Figure 10. Reliability of determinants: groupings analysis.

Figure 10 shows the values for rho^2 found for different grouping analyses. The figure indicates that regardless of the number of groups, building year determinants are most consistent, and the age of buildings within a grid-cell always scores the highest rho^2. The remaining four characteristics (FSI, GSI, number of floors and number of buildings) gain more consistency when 7 to 14 groups are created.

3.4.2 Identification of local and global salient urban areas

This section presents the results of step 5 on the identification of local and global salient urban areas. Regarding case study Amsterdam and parameter settings, as visualized in Figure 11 A, 494 local salient urban areas are distinguished, covering 11% of the built environment. Highest level local salient urban areas are represented by pink grid-cells and comprise, amongst others, Amsterdam Central Station and the Rijksmuseum. From the distribution within Figure 11 A, it may be expected local salient urban areas cluster more within the historical city center and many local salient urban areas are located near (intersections of) the bicycle street network. Subsequent analysis shows that neighborhood percentages indeed deviate from the city average, e.g., the historical city center has a coverage percentage of 16%, while prewar extension plans like Plan Zuid yield coverage percentages of 10%, whereas percentages for urban extensions during the 1960's, such as Westelijke Tuinsteden, are just above coverage 9%. The Gini coefficient is used to determine how local salient urban areas cluster near (intersections of) the bicycle network as will be elaborated on in 3.4.3.

Regarding case study Amsterdam and parameter settings, as visualized in Figure 11 B, there are 3.284 global salient urban areas covering 74% of the built environment. Highest level global salient urban areas are represented by bright pink grid-cells and are central locations, such as Dam Square, Damrak, and the Nieuwmarkt. The images A to C in Figure 12 indicate the historical city center, as a neighborhood, contains highest global salience (95% of the grid-cells have salience levels of 1 or higher). Just like the case regarding local salient urban areas, there seems to be a variation amongst urban expansion plans. For example, 70% of Westelijke Tuinsteden have statistically significant clusters of similar urban characteristics, while Plan Zuid reaches a coverage percentage of 58%. Furthermore, from the detailed Figure 12 B of Plan Zuid it can be seen that global salient urban areas follow the major axial streets.

3.4.3 Spatial distribution of salient urban areas

The last part of this section uses the Gini coefficient as a comparative measure of dispersion relative to salient urban areas within Amsterdam. This analysis is preferred over the multi-distance spatial cluster analysis because it is scale dependent (Tsai 2005). The ratio analyses are used to measure the inequality of the distribution of salient urban landmarks in Amsterdam, based on 1.) the extent to which an urban area is salient, and 2.) the number of salient urban areas within a certain distance field of a salient urban area. For example, a distance field of 300 meters represents 8 grid-cells surrounding a salient urban area. The Gini coefficient can range between 0 and 1, with 0 representing perfect equality, and 1 representing perfect inequality of the distribution of salient urban area in Amsterdam. Brown's formula has been used to calculate the Gini coefficients shown in Figure 13 A-B.

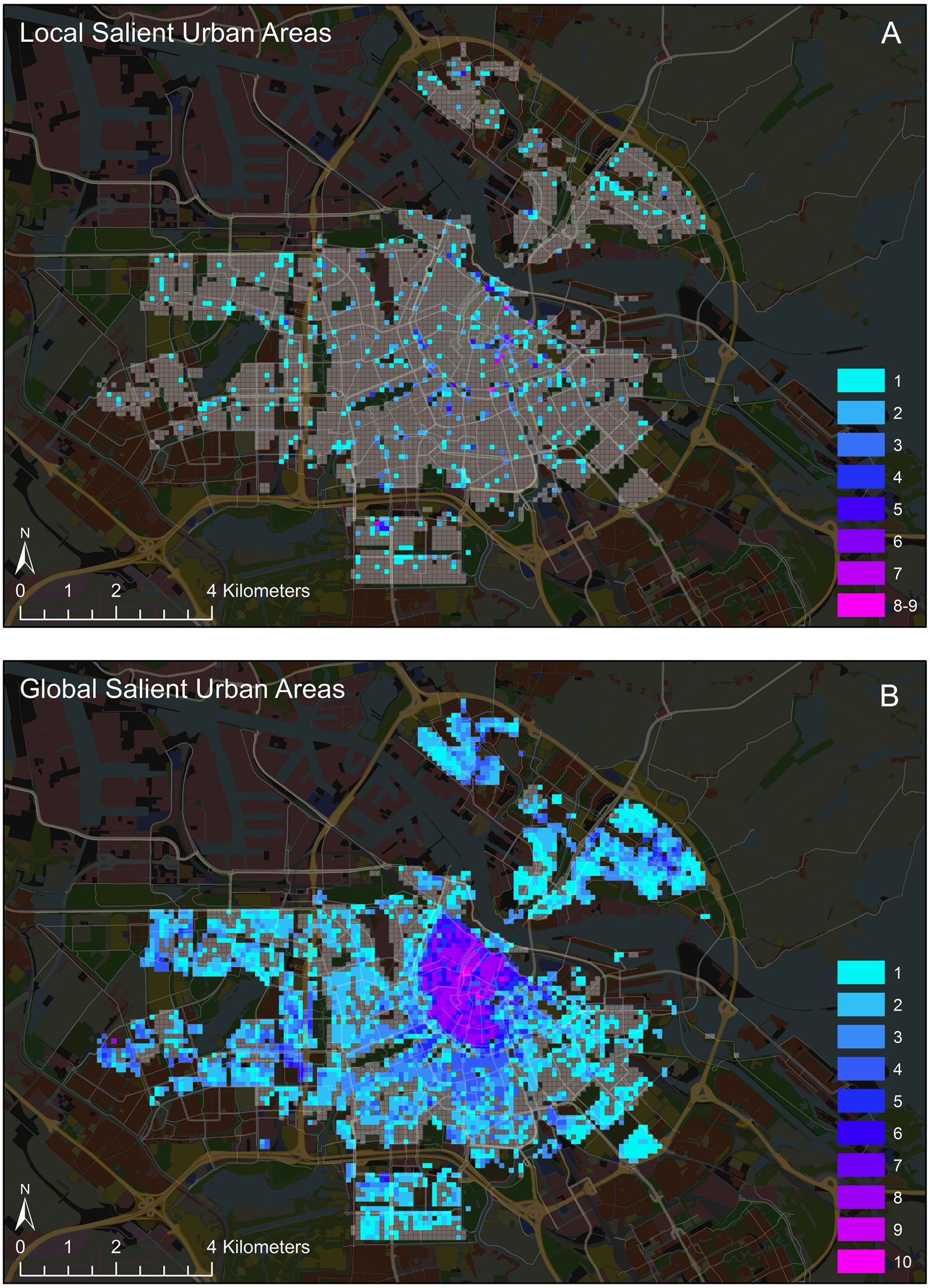


Figure 11 A-B. Identification of salient urban areas in Amsterdam.

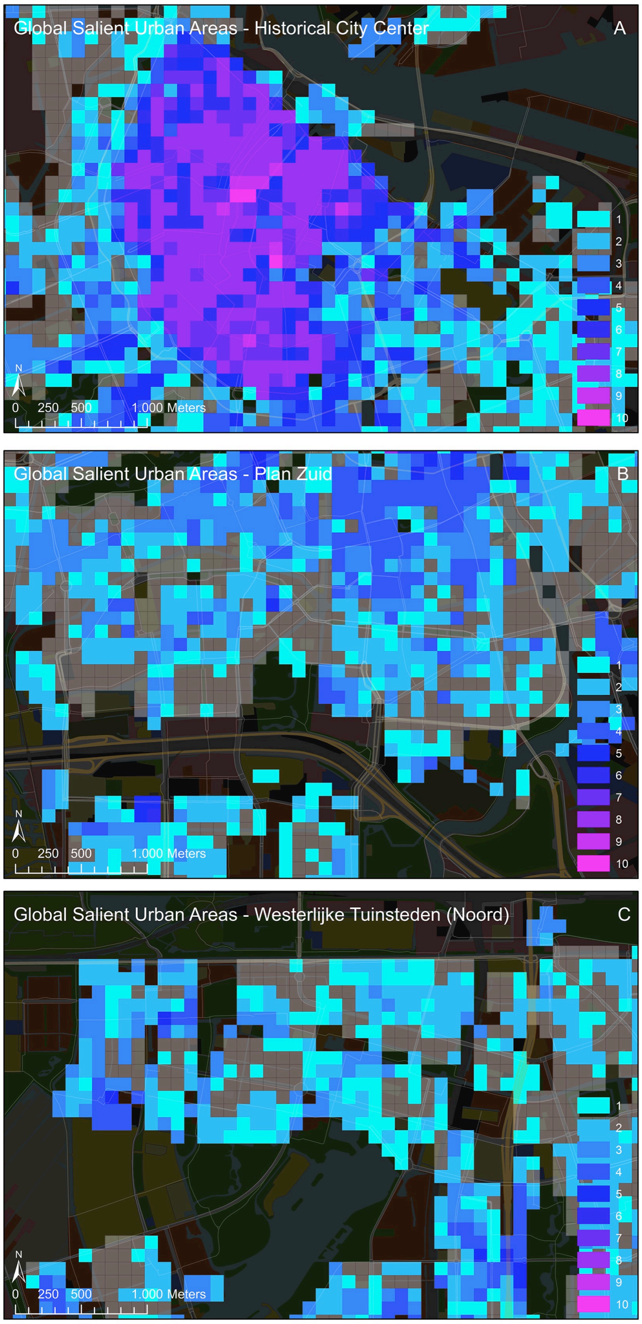


Figure 12 A-C. Detail images of neighborhoods and global salient urban areas.

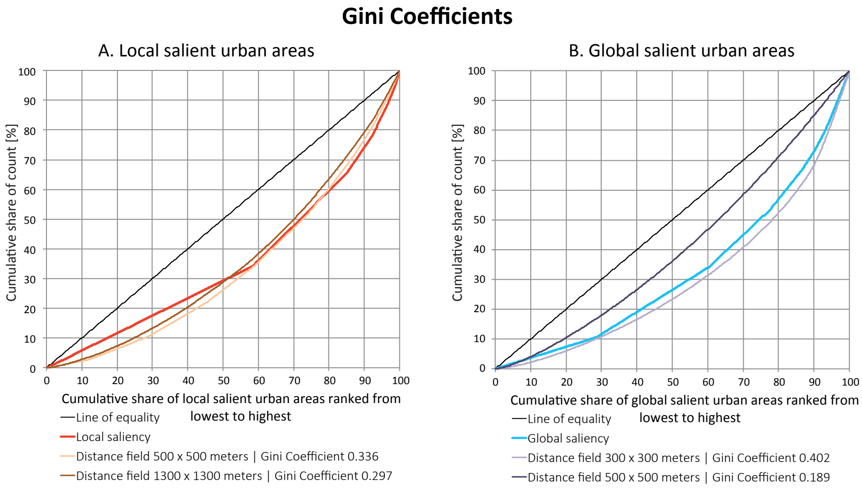


Figure 13 A-B. Gini coefficients of local and global salient urban areas.

The Gini-coefficient of saliency of local (and global) salient urban areas is 0,30 (0,35), meaning that saliency is distributed rather equally over all salient urban areas. Figure 13 A-B show that 58% of the local salient urban areas (28% of global salient urban areas) have only one salient determinant. These percentages correspond to 34% (local), and 10% (global) of the cumulative salience. Both Figure 13 A-B also indicate the 10% highest levels of salient urban areas correspond to 25% (local), and 29% (global), of the cumulative salience. The Gini coefficient representing the number of salient urban areas within a certain distance range fluctuates between 0,25 and 0,35 (local), and 0,19 to 0,23 (global), depending on the distance field.

In line with previous statements (3.4.2), this means saliency of local salient urban areas to be slightly more equally distributed compared to global salient urban areas. As to distance fields, more variation is found. The number of local salient urban areas within a distance field, is least equal at 500 square meters, and most equal at 1.700 square meters. On the other hand, the number of global salient urban areas, within a distance field, is most equal at 500 square meters, and least equal at 300 square meters.

In ArcGIS the number of local salient urban areas surrounding one local salient urban area can be visualized for different distance fields used to compute the Gini coefficient. Insights from these maps are complementary to the Gini coefficient, as the latter does not explain how salience is spatially distributed. Figure 14 A, for example, shows high values are concentrated around Vrije Universiteit van Amsterdam in the South, containing the smallest distance field of 300 square meters. Medium to high values are concentrated around larger public squares, such as, Central Station and Museumplein. Also, local salient urban areas with lower levels of salience appear to be located along major axial streets.

Figure 14 B shows spatial distribution changes according to different distance fields. E.g. a distance field of 500 square meters shows a concentration near Mr. Visserplein. Moreover, it becomes clear, more local salient urban landmarks with relative more local salient urban landmarks are distinguished within the proximity of 250 meters, such as around Vondelpark. By increasing distance fields, local salient urban areas within and bordering the historical city center gain higher percentages, meaning, it is more likely to encounter more local salient urban area when moving across the historical city center. Hence, routes across the historical city center are expected to be easier to memorize and structure in long-term memory.

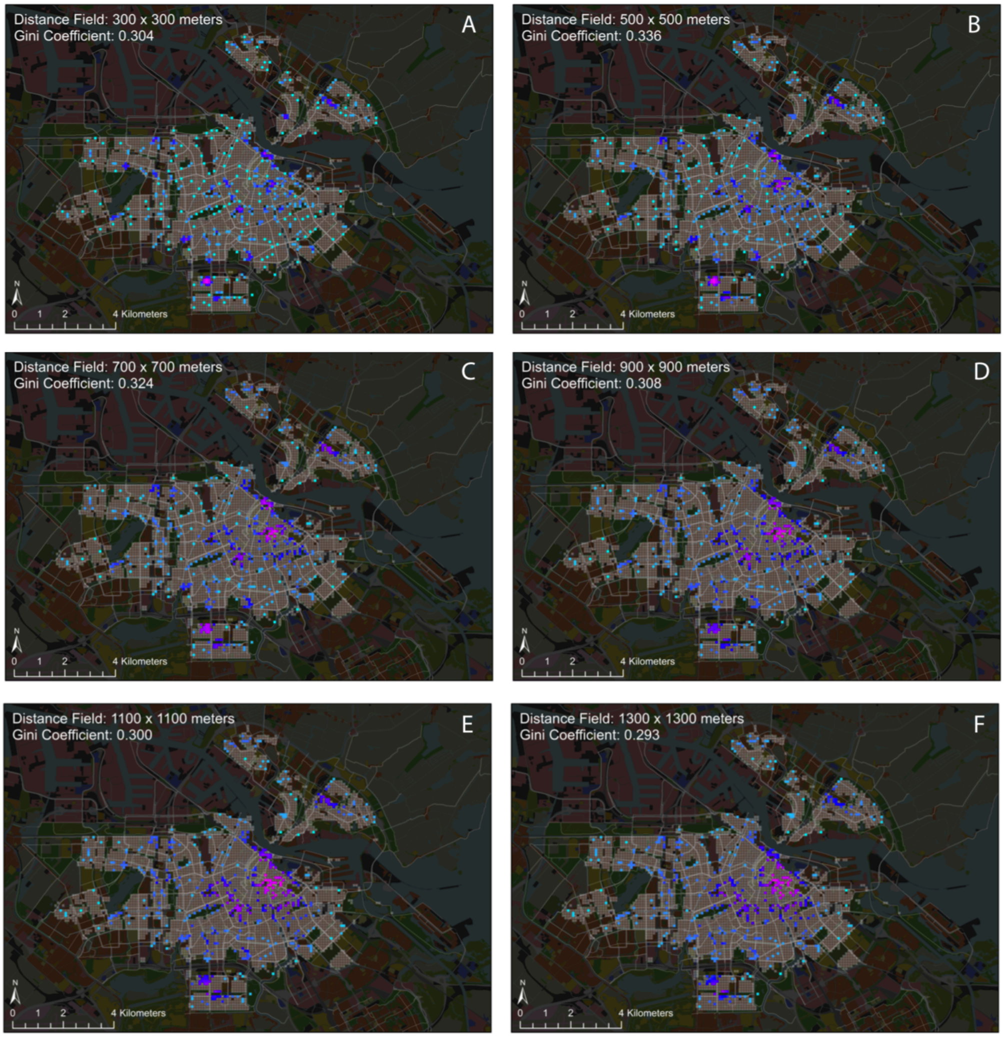


Figure 14 A-F. Spatial distribution of Gini coefficients for different distance fields.

3.5 Conclusion and recommendations

Landmarks are assumed to support wayfinding behavior in urban environments. Determining the location of distinctive landmarks is thus important for investigating route choice processes, structures of urban cognition, and travel information. However, currently most research approaches in this field require highly demanding data collection efforts. To overcome these demands, this chapter proposes an approach to handle open-source data.

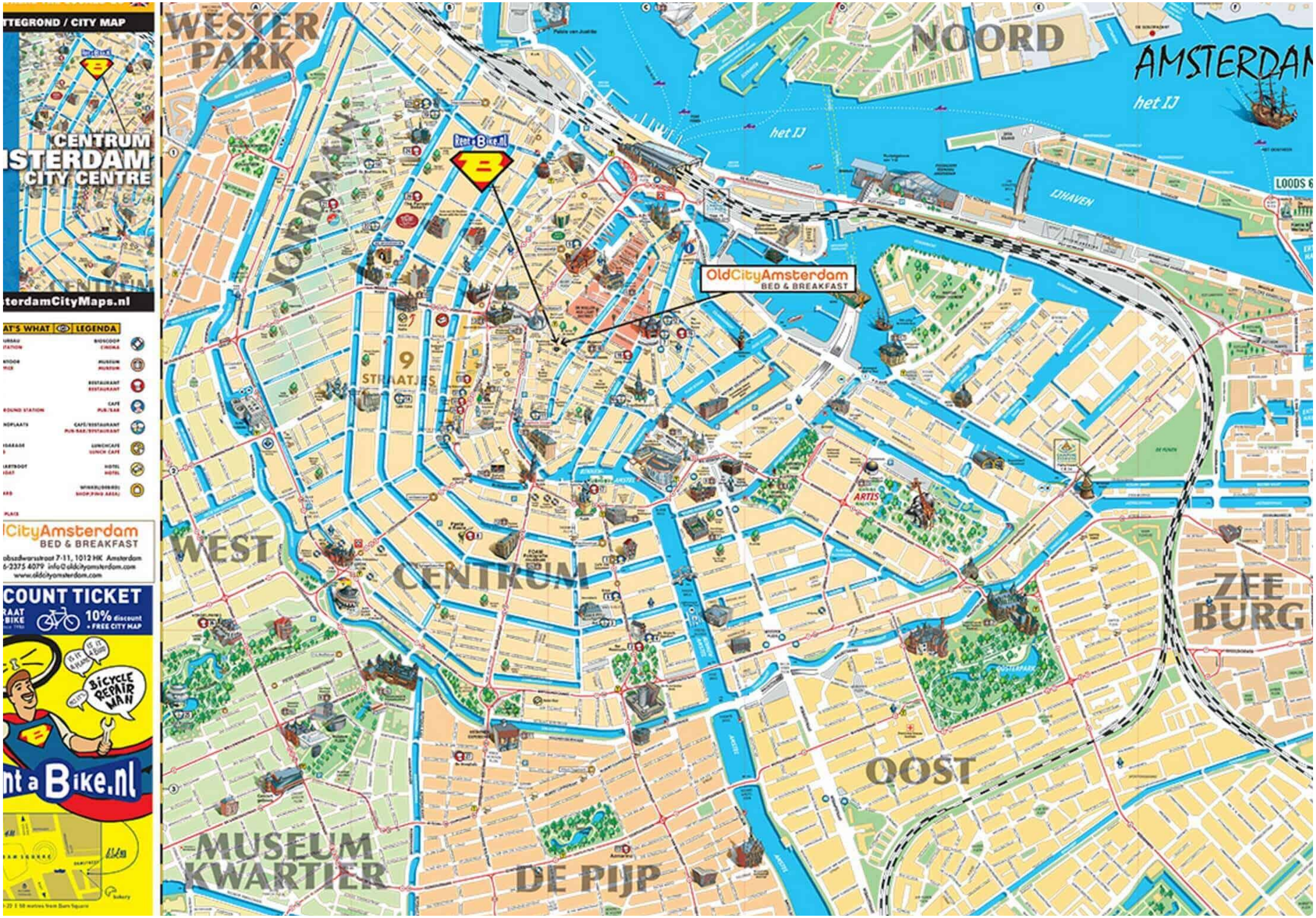
The proposed method combines insights from cognitive sciences and spatial analytics from urban morphologies to identify aggregated local and global urban landmarks based on salient characteristics. The method consists of five steps based on data management, grouping analysis, and cluster and outlier analysis. Results have been applied to identify the differences in distribution of cluster and dispersion between local and global salient urban areas using the Gini coefficient, based on an open-source GIS dataset on the built environment of Amsterdam.

Implications of identifying salient urban areas can provide new insights to analyze how wayfinding landmarks structure environmental knowledge and investigate influences on wayfinding strategies. This environmental knowledge (configuration of landmarks) is assumed to become available when also knowledge has been memorized about the general interrelationships between landmarks (Hirtle and Hudson 1991). If people use these wayfinding landmarks as part of the wayfinding strategy, this is expected to be observable in their route choice behavior. For example it could be more likely to take a detour if more wayfinding landmarks will be passed. Improved insights can potentially complement navigation apps, physical route signage, and urban planning.

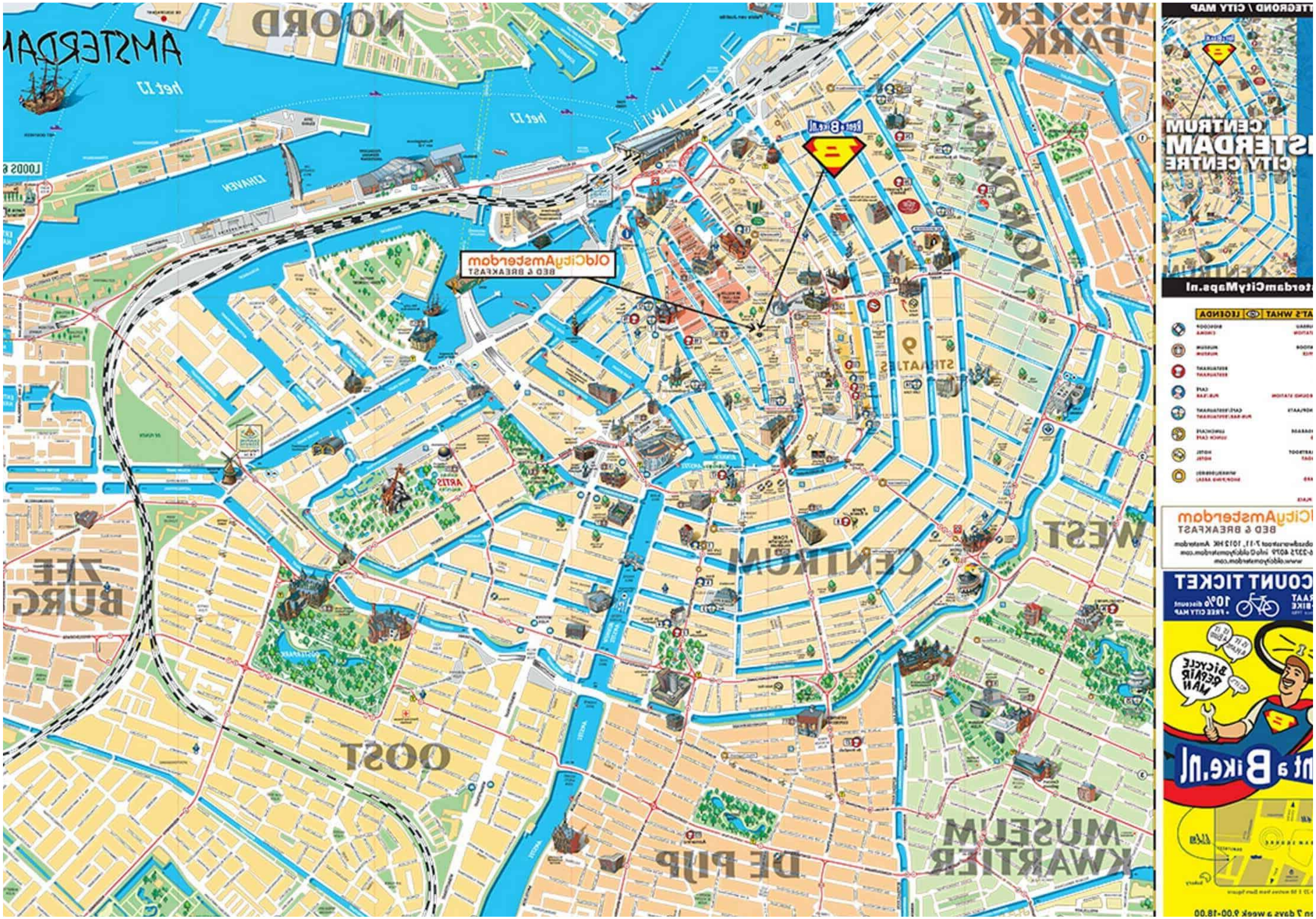
More research is needed to verify the parameter settings of this case study, and investigate other determinants. It is expected that digital elevation maps (AHN) or Lidar data will be a better indicator for building level. Further expansion of the determinants can also include traffic intensities, network characteristics, individual movement patterns using GPS and visibility using isovists, and functionalities. To improve validity it would be of interest to investigate to what extent the grouping analyses mimics the way people classify urban typologies through stated preference studies.

Acknowledgements

We thank the anonymous reviewers for the valuable comments and suggestions. The research leading to these results has received funding from the European Research Council under the European Union's Horizon 2020 Framework Programme for Research and Innovation. It is established by the Scientific Council of the ERC Grant Agreement n. 669792 (Allegro).



CENTRUM
STERDAM
CITY CENTRE
terdamCityMaps.nl
LEGENDA
lCityAmsterdam
BED & BREAKFAST
COUNT TICKET
10% discount
+ FREE CITY MAP
BICYCLE REPAIR MAN
nt a Bike.nl
WESTER PARK
NOORD
AMSTERDAM
het IJ
het IJ
IJHAVEN
LOODS 6
JORDAAN
Rent a Bike.nl
OldCityAmsterdam
BED & BREAKFAST
9 STRAATJES
ARTIS
WEST
CENTRUM
ZEE BURG
OOST
MUSEUM KWARTIER
DE PIJP


OldCityAmsterdam
BED & BREAKFAST
COUNT TICKET
10% discount
FREE CITY MAP
Bike.nl

4

Activity Patterns of Tourists in Amsterdam ft GPS

This chapter is an empirical study on tourist cyclists' mobility behavior in Amsterdam based on GPS tracks collected from bicycles that tourists have used. It uses multiple methods like clustering, different network analyses and activity spaces to characterize tourist behavior and activity zones of these cyclists and makes recommendations based on observed patterns. The results provide interesting findings about cycling tourists travel behavior and the structure of their cycling network and activity chains. Moreover, the processed GPS data can be used in subsequent studies, such as routing behaviour in Chapter 5.

1. Empirical study to unravel spatial and temporal characteristics of tourist activity patterns by bicycle in Amsterdam Metropolitan Region.
2. Activity detection algorithm to process GPS data into 10.347 and 105 activity locations and zones of 1.817 unique tourist day patterns.
3. Network analysis of spatial relations between activity zones indicates four tourism communities with similar behaviour to common transport networks.
4. New insights into spatial distribution of tourists based on relations between activity space, compacteness and travel time ratios.
5. Discussion on insights considers how tourists can be better spatially and temporally distributed.

This is an edited version of the following article:
Zomer, Duives, Cats, and Hoogendoorn (submitted). Activity Patterns of Tourists in Amsterdam ft GPS Bicycle Data.

4.1 Introduction

Amsterdam is a vibrant city, known for its canals, low-rise high-density urban environments, with the bicycle as the main mode of transport, and increasingly as a popular destination for tourists. The numbers for 2017 revealed a 13% increase in tourism in the Netherlands, of which nearly 50% remains in the Metropolitan Region of Amsterdam, against an average of 8% in Europe (UNWTO 2018a). In absolute terms there are almost 20 tourists per 100 inhabitants in the historical city centre of Amsterdam a day, less than Lisbon but more than Barcelona (Fedorova, De Graaff and Sleutjes 2018). UNWTO (United Nation World Tourism Organization) expects a continuous annual growth of (urban) tourism until 2030 led by prosperous economies and low long-haul transport costs. Tourism is important for people to exchange culture, and it is beneficial for (local) domestic product. Yet, "overtourism" leads to excessive noise, a nuisance for inhabitants, and pressure on infrastructure (UNTWO 2018b). This can create tension between citizens and tourists, and a decrease in the quality of life of citizens and the experience of tourists (UNTWO 2018b). Therefore, global strategies and measures to better understand and manage urban tourism, such as "disperse tourists within the city and beyond" have been established (UNTWO 2018b). However, the effect of the measures heavily depends on the travel behaviour of tourists within the respective cities.

An increasing number of tourists use the bicycle to explore the city and surrounding region of Amsterdam (Fietsersbond 2013; NOS 2015). A strong connection between the rural region, alternative urban activities, and main tourist attractions can be a sustainable, healthy, and inclusive opportunity to improve the quality of life and visiting experience. Notwithstanding, bicycle activity and movement patterns of tourists in metropolitan environments have received very little attention in science and practice. Our study aims to advance the understanding on tourist's activity behaviour by bicycle in a metropolitan region (Amsterdam) using GPS data and analysing these data using network and activity space analyses at both individual and aggregate levels. These insights into tourists' activity and movement patterns by bicycle and the evolvement of these patterns over time are important for retailers, policymakers and city planners to predict when and where these activities and associated movements are conducted. These insights are necessary to understand how urban travel behaviour of tourists can be influenced to i) stimulate new itineraries (activity sequences), ii) promote spatial dispersal, and iii) promote time-based dispersal.

The aim of this chapter is to determine the spatial and temporal travel and activity patterns of tourists with access to a bicycle in a metropolitan area. This chapter elaborates upon the design and execution of a large field study called LUCY, Learning and Understanding of CYclists behaviour. Longitudinal data are collected featuring the cycling behaviour of tourists in the Amsterdam Metropolitan Region. Measures for systematically analyze and quantify tourist activity patterns are developed and applied. This chapter provides empirical insights considering where, when, and for how long tourists perform activities while having access to a bicycle and identify the spatial relations between these activities. These insights can be used by policy-makers to improve urban planning and travel information to nudge tourists to less crowded areas.

The remainder of this chapter is organized as follows. Section 4.2 provides an overview of literature concerned with the data collection of tourists' movement behaviour and activity patterns, and research that provided insights into (general) cyclists' movement behaviour. Section 4.3 describes the data collection efforts of the LUCY project at The Student Hotel-City in Amsterdam. Section 4.4 presents the research methodology and discusses the key findings derived from an activity location detection algorithm and an activity space clustering analysis. The results of the analysis of the movement behaviour and activity patterns of tourists are reported in Section 4.5. Section 4.6 provides a discussion of the results and related conclusions regarding tourists' activity patterns.

4.2 Background

Bicycle mobility patterns and sustainable and inclusive tourism are receiving more attention from academia and society recently, but insights into urban travel behaviour of tourists by bicycle are limited. Therefore, the literature review features three topics. Section 4.2.1 elaborates on general travel behaviour in tourism research. Section 4.2.2 discusses research paradigms that have been used to investigate mobility patterns. Section 4.2.3. provides an overview of findings relating to bicycle tourists. Finally, in section 4.2.4 presents a synthesis and framework to position this chapter into the existing research realm of travel behaviour and tourism research.

4.2.1 Travel behaviour in tourism research

Tourism research ranges from qualitative studies based on literature (Larsen, Urry & Axhausen 2006) to quantitative studies based on travel patterns using GPS, mobile phones, or geotagged photos and tweets (Shoval & Raveh 2004; McKercher & Lau 2008; Pettersson & Zillinger 2011; Renso et al. 2013; Kádár & Gede 2013; Gong 2016). The understanding of tourist mobility has taken a leap in recent years, especially with the rise of GPS and mobile phone data (Shoval & Isaacson 2006). In short time, a strong body of knowledge has been gathered on behaviour of tourists walking in delineated or gated areas, such as zoos and natural parks (Hayllar and Griffin 2005; East et al. 2017; Meijles et al. 2014), and city centers (Shoval, Schvimer & Tamir 2018). Also, mobile phone data has proven to be an useful data collection technique to analyze large-scale travel patterns in a country or continent (Raun, Ahas & Tiru 2016). The majority of these studies focus on presenting empirical findings and discussing possible implementations. With respect to the mobility patterns of tourists, which have access to a bicycle, little research is presented.

4.2.2 (Tourism) mobility patterns

Another relevant topic relates to the prediction and modelling of tourist behaviour, mainly focussing on which activities will be performed where. Although discrete choice modelling allows studying the possibilities and limitation of information to divert visitors to alternative activity locations (Zomer et al. 2015), in tourism research most studies make use of other modelling paradigms, such as Markov Chains (Xia, Zeephongsekul & Arrowsmith 2009; Zheng, Huang, Li 2017).

Traditionally, transport scientists and geographers investigate the same issues with different definitions of mobility patterns. Whereas the latter usually contemplates on spatial distribution patterns or space-time diagrams, the former has a soft spot for economical models and optimization, including the discrete choice paradigm to investigate activity, mode, and route choices. From literature, it seems that network analyses provide a connection between both research areas (Haggett & Chorley 1969; Bell & Iida 1997; Sevstuk & Mekkonnen 2012; Zhong et al. 2015).

The first steps to quantify tourist's mobility patterns have been derived from travel diaries, clustering of activity location by different tourist typologies (Shoval & Raveh 2004), identification of experiences (Hayllar & Griffin 2005), and identification of eleven movement styles (McKercher & Lau 2008). Using time geography and travel diaries, differences in multi-day travel patterns of 73 tourists by car have been investigated in Sweden (Zillinger 2007). Since the rise of GPS, cellular, and public transport data, more recent studies on tourism mobility focus on individual and aggregate activity patterns (Ahas et al. 2007; Modsching et al. 2008; Shoval & Isaacson 2009; Pettersson and Zillinger 2011; Kádár and Gede 2013), distributions of hotels and activity locations (Shoval et al. 2011). However, new theoretical models that describe urban mobility patterns of tourists need yet to be established.

4.2.3 *The bicycle tourist*

There have been various research efforts to unravel commuter mobility patterns to improve travel demand models and inform policymakers. Topics of interest focus amongst others on how to include bicycle in mode choices (Ton et al. 2019), bicycle route choices (Broach, Dill & Gliebe 2012; Rasmussen et al. 2015; Ton et al. 2017), trip chaining behaviour (Schneider et al. under review), activity space of cyclists (Modsching et al. 2008; Schönfelder & Axhausen 2010). With more tourists exploring the city by bicycle, differences in mobility patterns between commuters and tourists become important.

However, empirical findings on urban mobility patterns of tourists travelling by bicycle are scarce. Yet, several qualitative studies provided behavioural insights into various concepts of bicycle tourism conducted among hard-core bicycle communities in Australia and the United Kingdom. For example, management during sport events (Buning & Gibson 2016); multi-day touring pattern (Ritchi & Hall 1999; Faulks, Ritchie & Fluker 2006), preferences and attitude (Ritchi, Tkaczynksi & Faulks 2010; Lamont & Causley 2010; Lee 2014), the theory of planned behaviour (Kaplan et al. 2015; Han, Meng & Kim 2017), sustainability (Lumdson 2000; Dickinson, Lumdson & Robbins 2011), and shared bicycles (Kaplan et al. 2015). Although the Netherlands has very good bicycle infrastructure and an active bicycle community, no research nor policy documents featuring mobility patterns of urban bicycle tourists exist to the best of the authors' knowledge. Thus, there remains plenty of ground to explore the activity and movement patterns of a growing group of urban bicycle tourists.

4.2.4 *Conclusions and research gaps*

In conclusion, limited research efforts have been reported on tourist behaviour by bicycle, especially in metropolitan regions. A recent study proposes a theoretic framework based on an extensive literature study (Caldeira & Kastenholz 2019). Furthermore, not many theoretical models and concepts have been developed or incorporated from daily travel behaviour (Miller 2003; González, Hildalgo & Barabási 2008; Song et al. 2010; Di Lorenzo et al. 2012; Schneider et al. 2013; Hasan et al. 2013) and systematically tested considering the large amount research on tourism mobility.

Our study aims to advance the understanding on tourist's activity behaviour by bicycle in a metropolitan region (Amsterdam) using GPS data and analysing these data using network and activity space analyses at both individual and aggregate levels.

4.3 Experimental design of LUCY

Revealed movement and activity behaviour patterns of people can be measured using mobile phones, GPS trackers, and geotagged photos. While sparse data allows identification of the spatial distribution, fine-grained data allows identification of trajectories and activity locations with great precision. Regardless of the sparsity level, several methodological steps are required to derive activity and mobility patterns from raw GPS data. Two months of GPS data from 250 bicycles of The Student Hotel Amsterdam-City (TSH) (www.thestudenthotel.com/amsterdam-city/) has been used to derive tourists' activity patterns during July and August 2017. During this period TSH houses only Hotel Guests (tourists and some students attending one of the summer schools at the University of Amsterdam). These tourists have the possibility to rent a bicycle during their stay for 9€ (or 12€) and voluntarily opt into the LUCY study at the beginning of the bicycle rental at the Student Hotel.

The GPS trackers are located under the saddle pin and charged by means of the power supply of the hub dynamo in the front wheel of the bicycle whenever the bicycle was moving. No personal or behavioural data has been collected to minimize burden and to protect their privacy. Although some tourists rent bicycles for multiple days, because of missing start and return times and occasional bicycle changes, each day every equipped bicycle receives a unique tourist id. With 542 day-rentals, most bicycles have been used during the last week of August, while the weekly average is 242.

Every 10 to 30 seconds the longitude, latitude, speed, and a timestamp are recorded for each bicycle. During July 1st 06:00:00 and September 1st 06:00:00 a total of 1.465.590 GPS points have been collected from 250 bicycles equipped with the trackers. Filtering and processing of the data are necessary to reduce inaccuracies and enrich the raw data to derive meaningful information on activity patterns of tourists featuring bicycles in Amsterdam. Heuristics have been developed as part of this process to identify each GPS point as part of an activity, a movement, or an outlier.

4.4 Data analysis approach

The raw GPS data requires substantial cleaning efforts due to fluctuations in time and space caused by low battery power, unfortunate characteristics of the built environment (e.g. high rise with glass facades), very low travel speeds, and historic cache memory of the tracker. Since the ground truth is absent, rules are devised to classify points as stationary (activity locations), moving or invalid prior to the deriving of activity and trip characteristics. In order to perform a meaningful analysis of the activity behaviour of tourists, a k-means analysis is used to assign activity locations to clusters (activity zones). The enriched data is then used to report aggregate statistics, and to perform network and activity space analyses. The complete data analysis approach is conducted in MATLAB and visualized in Figure 15.

4.4.1 *Preliminary data filtering*

The objective of the preliminary data filtering is to exclude extreme cases that may interfere with the state estimation process. The preliminary filtering is applied to a total of 1.465.590 GPS points that have been collected from 250 bicycles between July 1st 06:00 and September 1st 06:00. First, points that are visually not attached to an itinerary starting in Amsterdam have been removed from the database, totalling 642 points. Second, all GPS points with a registered speed and detected direction of 0 are excluded, as they represent points without speed and direction (1-360 degrees). Based on this criterion 76.985 additional points have been excluded. A total of 1.387.321 points were left after the preliminary data filtering.

4.4.2 *State estimation of GPS point*

The aim of the state estimation is to determine where and when activities are performed that required the tourists to park their bicycle for a reasonable amount of time. Whether or not an activity is performed depends on the state of the prospective GPS point and its neighbouring GPS points. A GPS point can be in one of three states, which can be described as stationary, moving, or invalid.

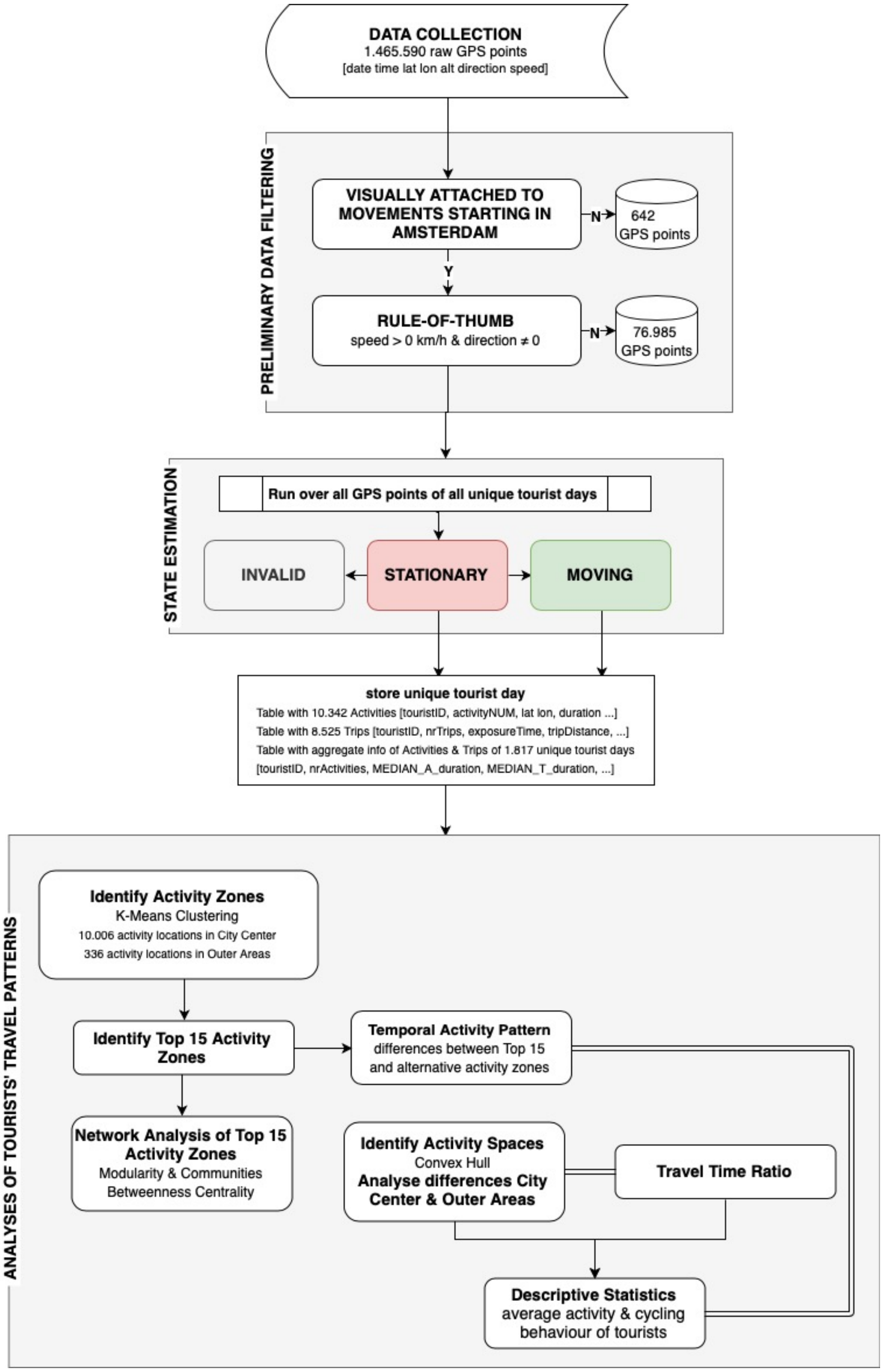


Figure 15. Four steps of the developed data analysis approach: preliminary data filtering, state estimation of GPS points, derivation of activity locations, and analyses.

4.4.2.1 *Identifying stationary GPS points*

Dependent on the duration, activities have an intrinsic value to tourists and derive a certain amount of satisfaction from participation in these activities. Therefore, there should be a minimum duration to identify activities. All stops for more than 3 minutes, the maximum cycle time of most intersection controllers in Amsterdam, could theoretically be an activity. As such, activities range from gathering travel information or taking a picture, to visiting a museum and going to a restaurant. A similar threshold can be found in various studies (Menghini et al. 2010) To alleviate instabilities when the GPS tracker is stationary, a minimum travel threshold of 400 meters and a maximum activity range of 150 meters are used. The travel threshold ensures two sequential activities are at least 400 meters apart, while the activity range ensures that also "very slow moving" can be detected as one activity. A visual sensitivity analysis at the individual level showed that these values yield the best results. In addition, the first and last GPS points are always stationary to ensure that a trajectory is always a complete travel chain. Therefore, the actual duration of the first and last activities on a given day cannot be inferred.

4.4.2.2 *Identifying moving GPS points*

Points in a moving state are part of a valid sequence of GPS points that are located between two activities that are at least 400 meters apart. Confidence to identify a point in moving state valid depends on the proximity and relative spatiotemporal position to previous and next points. There are no confidence issues when GPS points are fine-grained. However, a GPS point in a moving state is considered as scarce when the distance to the previous GPS point is more than 300 meters. Scarce GPS points typically occur due to low battery after a longer period of being stationary and signal loss while moving. During such periods the exact trajectory can only be estimated, and it is unknown if, and how long, the bicycle performed an (intermediate) activity. Based on a visual inspection of 300 tourist days more than 50% of the movement patterns have to be fine-grained in order to derive any meaningful insights.

4.4.2.3 *Identifying invalid GPS points*

GPS points can be identified as invalid in two ways, 1) if most of the movement pattern of a tourist day is incomplete, and 2) if there are unrealistic displacements within a tourist day.

- Incomplete activity and cycling behaviour at the individual tourist level. The aim of the study is to analyse activity-based behaviour of tourists, hence we require at least 3 activities (start and end are most likely at TSH, and at least 1 activity is visited in or around Amsterdam). Secondly, there is an expected minimum of 100 GPS points given an average cycling speed of 8 km/h, two trips of 525 ($\sqrt{2}$ * 400) meters, three activities with a minimum duration of 3 minutes, and a time interval of 10 seconds per point. Finally, the quality of the analysis will be affected if the number of GPS points within cycling movements is too often scarce (more than 50% of the travelled distance consists of scarce data points with more than 300 meters between "moving" GPS points).
- Remove spatial outliers. All points which result in unrealistic bicycling speeds are excluded, i.e. more than 40 km/h, and spatial outliers are identified as invalid points. The Hampel function is used to identify spatial outliers based on specified windows composed of a chosen number of neighbour GPS points (2 before, and 2 after the X and Y). Outliers in longitude and/or latitude are detected when both medians and both standard deviations using the median absolute deviation are higher than 2.

4.4.3 Derivation of activity locations

Upon excluding invalid GPS points, the classified stationary and moving GPS points that have been collected during July 1st and September 1st are processed and summarized for each tourist day, yielding 10.342 activity locations and 8.525 trips, made by 1.817 unique tourist days.

4.4.4 Analysing tourist activity patterns

Tourist activities form a network in space and time. To study their movements three analyses methodologies are adopted from complex network theory. A k-means clustering algorithm is used to identify main activity zones (4.4.4.1.). A network analysis is adopted to determine the spatial relations between activity zones based on the existence of communities and betweenness centrality indicators (4.4.4.2). Third, an activity space analysis to identify spatial differences between activity communities (4.4.4.3). The following sections will elaborate on the analyses and provide some preliminary results. A comprehensive discussion of the results can be found in sections 4.5.2 – 4.5.5.

4.4.4.1 Derivation of activity zones

The aim of this analysis is to derive activity zones that represent the main tourist destinations by bicycle in Greater Amsterdam Region. K-means clustering can be useful to an unsupervised algorithm to categorize and classify activity locations into activity zones based on the spatial proximity without a reference outcome. For privacy reasons, the bicycle is tracked instead of the tourist, which means that the final destination and actual activity type are unknown. Therefore, a spatial k-means clustering is performed on all 10.342 activity locations that indicate where tourists parked their bicycles. Generally, the Euclidean distance is used in spatial k-means clustering analysis. In this case, however, an Euclidean distance measure would provide unrealistic clustering results, given that most tourists by bicycle diverge from the direct (Euclidean) line between the identified activity location (of the parked bicycle) and the main destination (at the activity zone) due to street patterns in Amsterdam. A Manhattan (city block) distance computes the absolute differences between coordinates of pair of objects (Kaufman & Rousseeuw 2009), thus providing a more realistic clustering result. To ensure the avoidance of local minima 90 initializations are used.

The number of clusters is determined based on the minimum number of clusters where there exists a peak at mean silhouette value (i.e. the consistency of points within each cluster) compared to neighbouring clusters (Rousseeuw 1987) and the value of improvement of Best Total Sum of Distances, which should be higher than the average where the line stabilizes. However, activity locations are unevenly distributed among Amsterdam city centre and outer areas. Thus, two clustering procedures are performed, one for the locations inside the ring road and one for the locations outside the ring road. For this analysis 10.006 stationary data points are used from within the A10 highway surrounding the city centre, and 336 data points scattered in the outer areas. While a lower limit of the peak at mean silhouette value for the city prevents an overestimation of activity zones in known and main locations in Amsterdam, an upper limit for the outer areas prevents an underestimation of activity zones (e.g. combining Haarlem, Bloemendaal aan Zee, and Zandvoort into one cluster).

The clustering analysis resulted in 60 activity zones (i.e. clusters) within the city centre of Amsterdam and 45 activity zones in the outer areas of Amsterdam (depicted in Figure 16). In the remainder of this chapter, a combination of both sets of is used, totalling 105 activity zones, to analyse the activities, trips and the aggregation at the individual tourist level. For each day tourist, all activity locations are linked to an activity zone, and the sequence of visited activity zones (activity sequence) has been detected. In addition, the activity zones are used to identify the spatial relation between activity zones based on the existence of communities and betweenness, and (iii) to analyse differences in activity spaces.

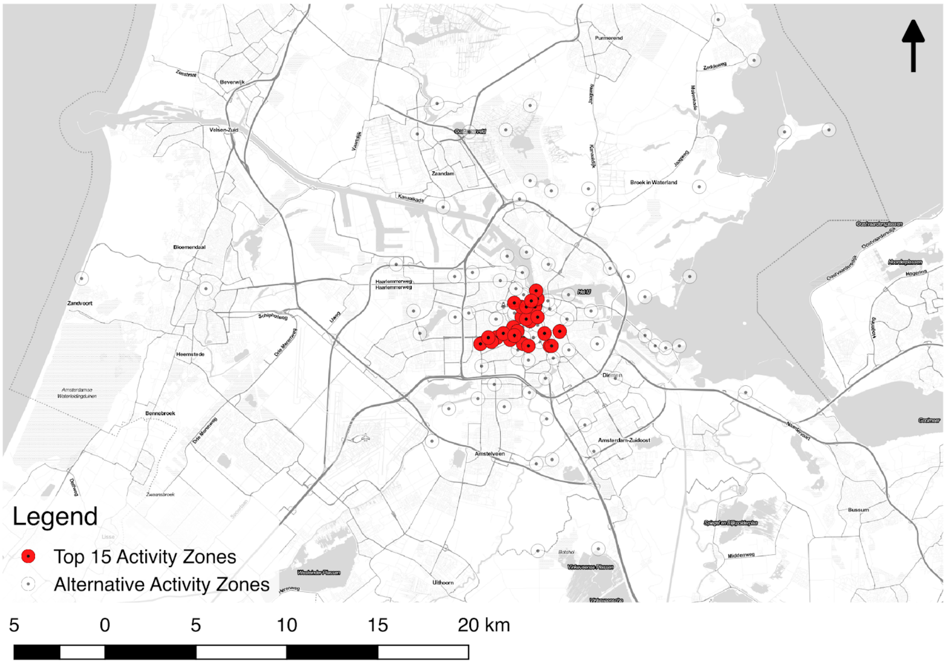


Figure 16. Spatial distribution of the identified activity zones, where the Top 15 activity zones are indicated with a red colour.

4.4.4.2 Network analysis: communities of activity zones

The network analysis aims to reveal the topological features to understand the dynamics of the activity zones network resulting from the activity pattern. Of interest is to investigate if there are so-called *communities*, i.e. set(s) of activity zones that are generally visited in combination on a given day by tourists of TSH. For instance, if the majority of the tourists that visit(ed) Museum Square also visit(ed) the Vondelpark followed, or preceded, by Leidse Square on the same day the three activity zones are likely to belong to the same community. The existence and composition of communities will influence where wayfinding systems, and which content, should be located to a) stimulate tourists to remain in a specific community, and b) distribute tourists to other, less crowded, communities. Differences can be observed in the characteristics of communities in terms of weighted degree, clustering coefficient, and betweenness centrality (Newman 2006).

First, the number of communities in the tourists' activity zone network can be derived, based on the maximal *modularity*. The modularity computes how many communities exist in a network. That is how many communities have more trips to activity zones within the community than expected by chance, and fewer with other communities. To investigate activity patterns of the most crowded areas in Amsterdam, in this study the optimal number of communities is based on the modularity M of the Top 15 most attractive activity zones that amount to 74% of all activities, and their corresponding 7 most visited destinations. The selected activity zones are in line with suggestions several online tourist guides. As some of the destinations are not in the Top 15, the total number of nodes N is 28, with 105 links L. If the number of links (L_c) within a community c is larger than the expected number of links between (N_c) nodes of that community c given the networks' degree sequence, then the nodes of sub-graph (C_c) could, in theory, be part of a community (Blondel et al. 2008; Barabasi 2016).

$$M = \sum_{c=1}^{n_c} \left[\frac{L_c}{L} - \left(\frac{k_c}{2L} \right)^2 \right] \quad (2)$$

Where K_c is the sum of degree centrality for all the nodes in community c.

Based on the Top 15 activity zones and their corresponding 7 most attractive destinations this approach identifies 4 communities with a modularity of 0,117.

4.4.4.3 Network analysis: betweenness centrality

The consequence of this derivation of modularity is that for an activity zone to be part of a community, it has to be visited in combination with two or more activity zones of that community. Thus, tourists that visit one activity zone of a community are likely to also visit other activity zones of that community. At the same time, transport networks generally have both strong ties between communities and weaker ties within the community. Therefore, it is hypothesized that the number of connections between activity zones correlates with the betweenness centrality (Barabasi 2016). *Betweenness centrality* is defined by the sum of the number of all shortest paths that pass through an activity zone (*v*) between the total number of shortest paths between each activity zone (*i*) and other activity zones (*j*) in the network, while omitting loops (Brandes 2001).

$$C_B(i) = \sum_{i \neq v \neq j \in V} \frac{\sigma_{ij}(v)}{\sigma_{ij}} \quad (3)$$

Three network characteristics are expected to correlate with betweenness centrality of activity zones (Newman 2006), if it indeed functions like a transport network:

1. *Closeness centrality* is calculated as the sum of the distance (*d*) of all shortest paths between each activity zone (*x*) and other activity zones (*y*) in the network (Bavelas 1950). The higher the value, the smaller the proximity to other activity zone.

$$C_C(x) = \frac{1}{\sum_{j, i \neq j \in V} d(y, x)} \quad (4)$$

2. *Weighted degree* is calculated as the sum of the weights (w_{ij}) of tourist bicycle flows between activity zones *i* and *j*.

$$s_i = \sum_{j \in \Pi(i)} w_{ij} \quad (5)$$

3. *Clustering coefficient* (C_c)is calculated based on the average local clustering: the probability that adjacent activity zones (*k*) are connected with each other based on the number of triangles (*t*) incident to the activity zone (Watts & Strogatz 1998; Newman 2000).

$$c_c(i) = \frac{2t}{k(k-1)} \quad (6)$$

Three hypotheses can be formulated, given the three characteristics of betweenness centrality and communities, and the assumption that tourists' activity patterns are behaviourally similar to other transport networks:

- Betweenness centrality has a positive relation with the closeness centrality and weighted degree;
- Betweenness centrality has a negative relation with the clustering coefficient, and;
- There are differences between the four communities with respect to weighted degree, closeness centrality, betweenness centrality, and clustering coefficient.

4.4.4.4 Revealed activity space analysis of tourists featuring bicycles

The aim of the activity space analysis is to identify spatial differences between activity communities. Activity space depicts the area where activities are performed by an individual (Newsome, Walcott & Smith 1998). The activity space of commute behaviour is often based on activity chains (primary activity, i.e. home - secondary activity, i.e. grocery shopping/pick up - primary activity, i.e. work/education, … - primary activity, i.e. home) and used to identify the area where activities are likely to be performed considering time and spatial constraints. Tourists activity behaviour is presumed less hierarchical compared to commuters because mandatory, preplanned activities, such as work or education, are rare. In this chapter activity locations of each tourist day have been used to determine the *revealed activity space*. Therefore, the convex hull is used to compute the Euclidean space surrounding the activity locations a tourist chooses to visit on a given day. The aim is to analyse if the revealed activity space and corresponding activity pattern are significantly different depending on visiting activity zones within the city (Central Station) or in the outer areas of Amsterdam (Zaandam Region). Revealed activity spaces are described based on the following 4 metrics:

4. The area of the convex hull depicts the size of the revealed activity space.
5. The perimeter is the minimum distance needed to enclose the convex hull (Manaugh & El-Geneidy 2012).
6. Compactness is the ratio between the area of the convex hull and the square of the perimeter, times 4π to ensure values between 0 and 1 (Selkirk 1982; Manaugh & El-Geneidy 2012).
7. The travel time ratio is derived by dividing the median travel time of a tourist by the sum of the median travel time and median activity duration for the same tourist (Dijst & Vidakovic 2000).

It is hypothesized that the revealed activity space is constrained due to a trade-off between travel time and activity duration.

The 1.817 revealed activity spaces cover on average an area of 3,6 km², with a perimeter of 9 kilometres, resulting in a compactness of 0,35. This indicates that on average tourists visit one activity zone located 4,5 kilometres from the hotel. The average travel time ratio of the tourists is 0,67, hence the average ratio between activity and trip duration of tourists is 1 : 3,2.

4.5 Four W's, One H: Revealing activity patterns of urban bicycle tourists

This section provides insights into the activity patterns of tourists that have access to a bicycle rented from a hotel in the city centre. Our aim is to determine where and when tourists perform activities and whether there are relations between activity locations. This section first elaborates on who the bicycle tourists are (4.5.1). Accordingly, the activities locations visited by these tourists are discussed (4.5.2). Third, the relation between the derived activity zones is determined (4.5.3). Afterward, this section analyses the activity space of tourists (4.5.4). This section finishes with an analysis of the time at which tourists bicycle trips are performed in Amsterdam (4.5.5).

4.5.1 Who are the cycling tourists?

During the holiday season, Hotel Guests of The Student Hotel-City (TSH) come from all over the world, the most common countries of origin of tourists renting bicycles are Germany, France, and Belgium. Hotel Guests can rent a Van Moof bicycle during their stay for 12 euros per day (including insurance). Although the average duration of stay stretches over multiple days, the stereotypical Hotel Guest rents for two people a bicycle for only one day. Therefore, it may be assumed that to some extent these tourists have more or less planned this day to travel around by bicycle, and they used other travel modes during the other days. For the anonymization of the data, all movements made by one tracker on a single day have a unique tourist day. This means that tourists that rented a bicycle(s) for more than one day have been assigned to multiple unique tourist id's. The actual total number of rented bicycles during the data collection period is 1.936. However, 8% of the actual unique tourist days have been excluded during data cleaning and processing. The remaining analyses pertain to 1.817 unique tourist id's that at least performed 3 activities, with high-quality GPS data (more than 100 GPS points, and more than 50% of the trajectories consist of fine-grained GPS points). A total of 8.525 bicycle trips and 10.342 activities have been identified (see Table 10). On average tourists are active (cycling and performing activities in the city) for 4 hours and 43 minutes with a total cycling distance of 15,14 km on a day.

Table 10. Descriptive statistics of tourist travel behaviour according to the LUCY study (July-August 2017).

Attribute	**Tourists in Amsterdam**		
# Individuals	1.817*		
# Activities	10.342		
# Bicycle Trips	8.525		
*Individual**	**Range**	**Mean**	**Std.**
# Activities	3-30	5,69	2,51
Activity duration [min]	0h - 6h29m	15m	21,4
Trip distance [km]	0,03 – 24,32	3,16	1,93
Bicycle duration [min]	0h1m - 11h13m	48m	70
Bicycle speed [km/h]	0,13 km/h - 84., km/h	8 km/h	4,90
Travel time ratio	0,01 - 1	0,67	0,25

**based on median of individual travel behaviour*

During the same period, “Onderzoek Verplaatsingen in Nederland” (OViN) collected data on the mobility patterns of citizens of the Province Noord-Holland. The aggregate results of OviN indicate that, on average, citizens make 0,71 trips by bicycle in July and August 2017 (CBS 2019). It should be noted the average number of trips for citizens is 2,34 when all travel modes are considered (CBS 2019). The average trip distance of a citizen cycling trip is 3,08 km, with a travel time of 15,21 minutes, yielding an average bicycle speed of 12,15 km/h (CBS 2019). Table 10 shows that behaviour of tourists participating in the LUCY study. Compared to locals (commuters and/or citizens), tourists renting a bicycle at TSH perform more trips; they visit on average 5,69 activities, with an average activity duration of 15 minutes. The trip distance is similar (on average 3,16 km), but the travel time is more than 3 times longer and, as such, tourists have a far lower bicycle speed.

From literature travel time ratios of citizens are known by trip purpose, e.g. 0,11 for a commute, 0,40 for shopping trips, and 0,25 for social leisure trips (Schwanen & Dijst 2002; Susilo & Dijst 2009). Although this study does not contain information on trip purpose, an average travel time ratio of 0,67 reveals that there is a difference between activity and trip duration of tourists and citizens, 1 : 3,2 for tourists versus 10 to 2,5 : 1 for citizens. Thus, tourists spend on average three times more time travelling by bicycle than performing the activity, while citizens spend 10 to 2,5 times more time at the activity (respectively commute and shopping) than travelling. This finding indicates that bicycling is not only a means of transportation; it is also the main activity of the day. It should be noted that aggregation at the individual level yields a higher travel time ratio compared to the activity-trip based average in Section 4.5.6, where the ratio appears to be closer to 0,50 (1 : 1). This indicates that the aggregate descriptive at the individual level is slightly biased towards longer bicycle trips and/or shorter activities. Regarding travel time ratios for citizens from literature, the averages are not solely based on bicycle movements, also car and public transport is included.

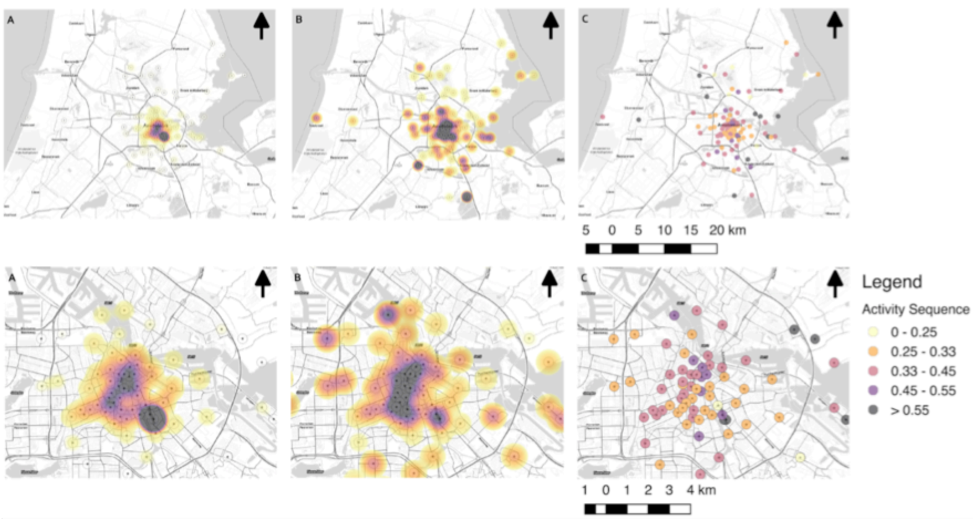


Figure 17 A-E. Overview of activity zones A) heatmap weighted by the number of visits in each activity zone, B) heatmap weighted by the average activity duration in each activity zone, C) activity sequence indicates which activity zones are on average visited at the beginning (0 – 0,5) or end (> 0,55) of the activity pattern.

Table 11. Characteristics of the Top 15 most visited activity zones.

#	Size	Name	Duration (mean)	Start time	Most common next destinations	Rel. activity [0 - 1]
1	3.464	The Student Hotel (*combined*)	53 min	00:00 16:30	• Albert Cuyp Market • Waterloo Square & Hortus Botanicus • HvA & UvA City Campus	0.,58
2	463	Central Station (*combined*) (transport hub, canal cruises, ferry to Amsterdam-North)	46 min	16:03	• TSH • New Market • Canals / Brouwersgracht	0,47
3	459	Museum Square (*combined*) (RijksMuseum, Van Gogh Museum, P.C. Hooftstraat)	54 min	14:26	• TSH • Vondelpark • Leidse Square	0,32
4	454	Albert Cuyp Market (*combined*) (Sarphati Park, Frederik Square)	36 min	15:44	• TSH • Museum Square • Vondelpark	0,36
5	429	Vondelpark (*combined*)	25 min	15:37	• TSH • Leidse Square • New Market	0,37
6	361	New Market (*combined*) (Red Light District & China Town)	32 min	16:30	• TSH • Dam Square & Rokin • Canals / Leidsestraat	0,48
7	297	Waterloo Square & Hortus Botanicus	24 min	14:02	• TSH • New Market • Dam Square & Rokin	0,29
8	288	Leidse Square	57 min	14:50	• TSH • Vondelpark • Munt Square	0,4
9	222	HvA & UvA City Campus	18 min	13:33	• TSH • Albert Cuyp Market • Museum Square	0
10	216	Rembrandt Square	31 min	16:56	• TSH • Munt Square & Flower Market • Dam Square & Rokin	0,5
11	212	Munt Square & Flower market	57 min	15:07	• TSH • Brewery ‘t IJ • Spui	0,33
12	210	Spui (shopping district)	46 min	15:33	• TSH • Vondelpark • Canals / Leidsestraat	0,43
13	199	Dam Square & Rokin (shopping district)	51 min	16:03	• TSH • Munt Square & Flower Market • Waterloo Square & Hortus	0,4
14	197	Anne Frank Museum & Jordaan	46 min	15:04	• TSH • Rembrandt Square • Waterloo Square & Hortus	0,38
15	192	Oosterpark	27 min	14:31	• TSH • Park Frankendael • Brewery t’IJ	0,33

4.5.2 *What do tourists visit by bicycle?*

The distribution of tourists is studied using 105 activity zones, i.e. clusters of activity locations derived through the k-means analysis. Figure 17 A-C show the distribution of activity locations over the 105 activity zones (i.e. degree) (A), the average activity duration (B), and activity sequence (C). The largest cluster size contains 2.549 activities at The Student Hotel (TSH), while the smallest clusters contain a single activity location. In a few occasions, multiple clusters are identified at several Point-Of-Interest locations (e.g. 3 near TSH, 4 within the Vondelpark). In the remainder of the study, these clusters are combined into one activity zone. A preliminary analysis shows that the maximum and minimum average duration at an activity is 337 minutes, and 2 minutes and 47 seconds, respectively. Activity zones that typically visited first have an average relative activity sequence lower than 0,33, while activity zones visited more often at the end of the tourist day correspond to values above 0,5.

Table 11 lists the 15 most frequently visited activity zones which amount to 74% of all activities. Noticeably, all these clusters are located within the city center boundaries. The maximum average duration of the Top 15 activity is 57 minutes (Leidse Square and Munt Square & Flower Market), and the shortest is 18 minutes (HvA & UvA University Campus). Activity zones that are generally visited at the beginning of the tourist day are located within a 3 kilometer radius from TSH (e.g. HvA & UvA University Campus, Waterloo Square & Hortus Botanicus, Museum Square, Munt Square & Flower Market, and Oosterpark). Activity zones that are generally visited on “the way back” are located further away from the TSH (e.g. Central Station, New Market) and more often located in the outer areas, outside the A10 beltway (e.g. Monnickendam, Johan Cruijf ArenA).

4.5.3 *What are the relations between activity zones?*

In general, tourists visit multiple activity zones for one day. This section studies the flow of tourists between activity zones. Activity communities were used to determine the activity zones that are often visited in a day within a given day.

Figure 18 illustrates the spatial relations between the Top 15 most visited activity zones and their 7 most frequent destinations. Four tourist communities are derived from the modularity (Blondel et al. 2008). Based on the urban environment and building functions the 4 communities can be coined as the Historical City Centre - HCC (purple), Party Places - PP (orange), Vondelpark & Museum District - VMD (green), and Most Popular Alternative - MPA (blue). Visually, the communities seem to be located in rather distinct areas. Yet, the modularity value indicates that the communities are weak (0,117), which implies that on a day tourists visit activity zones that are part of multiple communities.

Based on Table 11 and considering tourists visit on average 5,67 (6) activity locations, the most likely activity zone chain are depicted in Figure 19. The figure demonstrates that there are two major activity chain typologies, 1) upon visiting the Albert Cuyp Market activity zone, the majority of the sequential activity zones are located in Vondelpark & Museum District, and 2) upon visiting the Waterloo Square & Hortus Botanicus activity zone, the majority of the sequential activity zones are located in the Historical City Centre. The top activity chain possibly belongs to summer school students who start the day with lectures at the city campus, and in the afternoon following a similar activity sequence as other tourists. The figure also indicates that tourists have a rough activity schedule from the start as both Albert Cuyp Market and Waterloo Square & Hortus Botanicus are visited at the beginning of the activity chain.

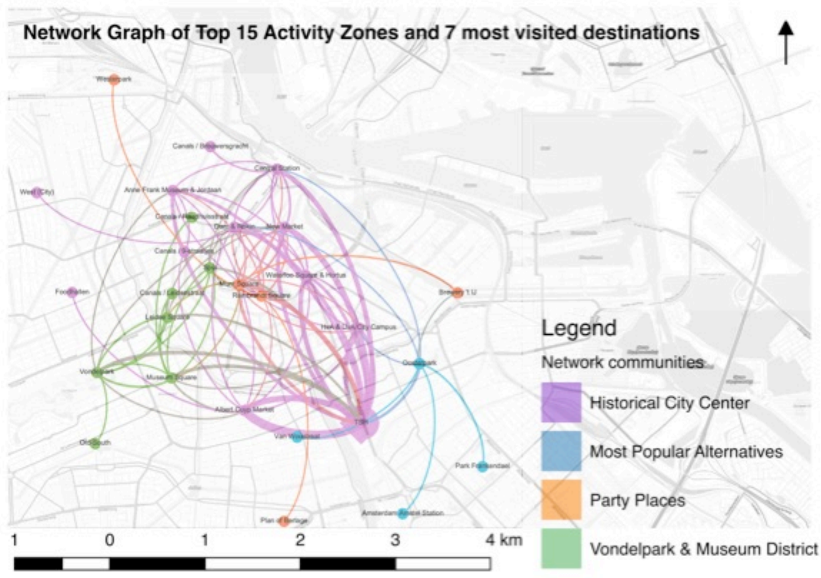


Figure 18. Gephi visualisation of the tourist network where the colors reflect the tourist communities based on the modularity (purple: Historical City Centre, green: Vondelpark & Museum District, orange: Party Places, and blue: Most Popular Alternatives) and the width of the lines the weight.

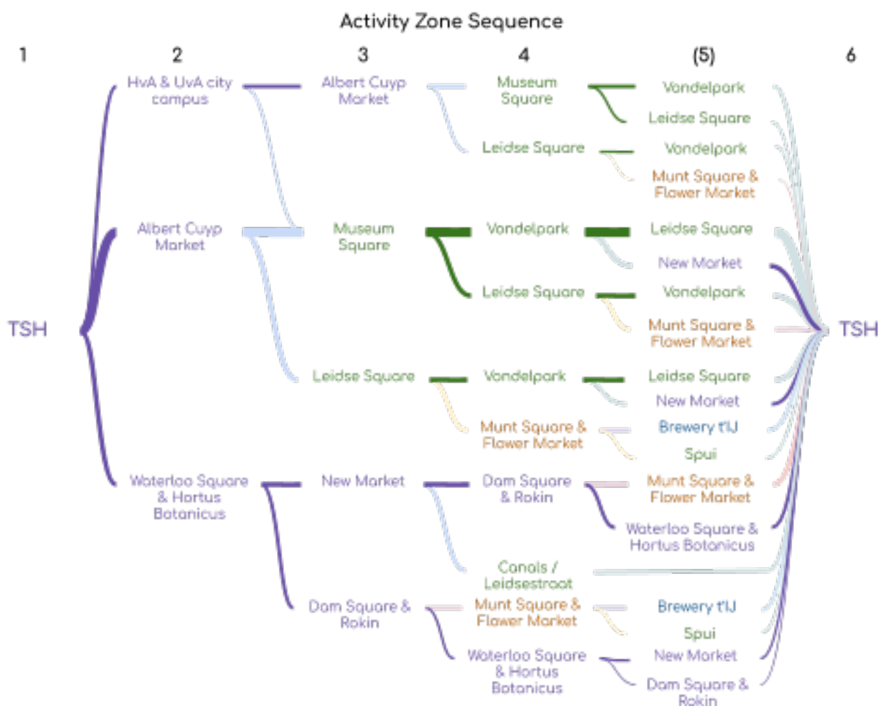


Figure 19. Most likely activity chains (colour/sign indicates corresponding community).

Figure 20 A-F relate betweenness centrality with the closeness centrality, weighted degree, and clustering coefficient of the 15 most visited activity zones per modularity. The data indicates minor differences among the identified communities. On average, Part Places activity zones have the highest betweenness and closeness centrality, Historical City Centre activity zones have a higher weighted degree, but Vondelpark & Museum District activity zones have, a higher clustering coefficient.

The Dam & Rokin activity zone has the highest betweenness centrality. This finding suggests that Dam & Rokin is an activity zone that functions as a bridge between other activity zones. Other dominant "bridges" in the tourist network are Munt Square and Rembrandt Square. The Student Hotel has a significantly higher degree than any of the other activity zones, because tourists "come back" to the hotel from 50% of the activity zones.

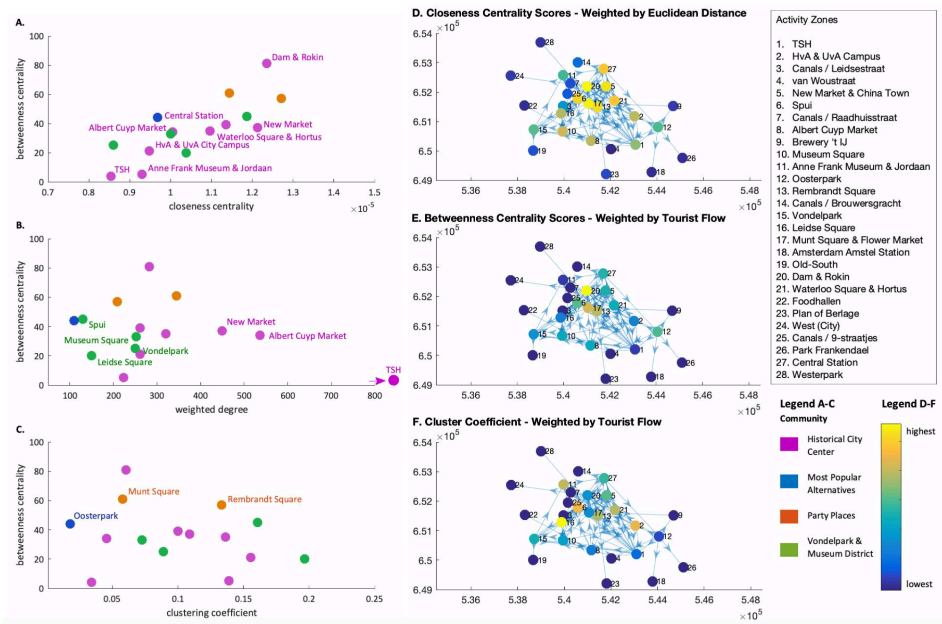


Figure 20 A-F. Network characteristics correlation with the betweenness centrality of Top 15 activity zones, correlations with A. closeness centrality, B. weighted degree, C. clustering coefficient, spatial distributions of D. closeness centrality, E. betweenness centrality, and F. cluster coefficient.

From literature, it has been hypothesized that betweenness centrality has a positive correlation with the closeness centrality and weighted degree. Figure 20 A shows that closeness centrality exercises a positive correlation with betweenness centrality. This implies that there is variation in the activity pattern of tourists, even among the 15 most visited activity zones. Three highly connected activity zones also have the most central locations and can be reached without much effort (Dam & Rokin, Munt Square & Flower Market, and Rembrandt Square). These activity zones are important to guide tourists to less crowded areas in the city.

Regarding the weighted degree, the data shows two trends, (i) the lowest values have a positive relation with betweenness centrality, while (ii) higher values show a negative relation with betweenness centrality (Figure 20 B). The three outliers are most likely the result of the start and end locations of the unique tourist days. Most of the times the trajectory starts and ends at The Student Hotel, but sometimes the trajectories start only after the tourists already travelled due to charging of the battery. Also, it has been observed that occasionally the bicycles stay a night in the city centre. Another explanation could be that the location of the hotel influences the travel behaviour of tourists; due to the proximity of (and promotion by) the TSH exceptionally more tourists visit Albert Cuyp Market and New Market by bicycle. This would mean that the location of (and promotion at) hotels can be used to distribute tourists to less crowded areas. The other Top 15 activity zones are in line with the hypothesis, and it proofs that the network resulting from tourists activity pattern behaviour is similar to common transport networks, i.e. activities have strong ties between communities and weak ties within the community.

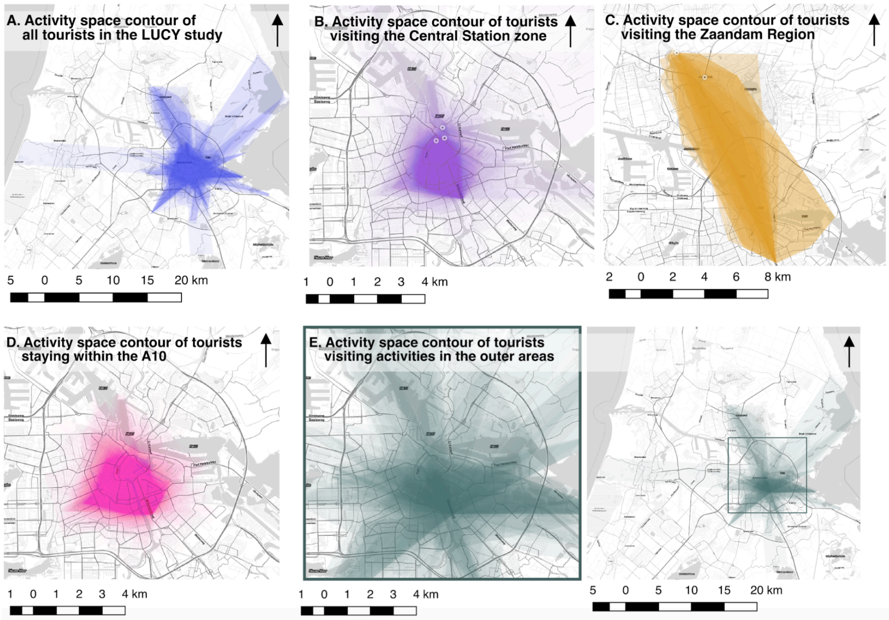
In addition, it has been hypothesized that betweenness centrality has a negative relation with the clustering coefficient in most transport networks. Figure 20 C shows indeed a negative linear relation between the clustering coefficient and betweenness centrality of the tourists' activity network by bicycle. Spatially adjacent activity zones with high betweenness centrality within the same community are generally not connected. This finding implies that tourists visiting an activity zone with low betweenness centrality are likely to visit other activity zones in the same community (Waterloo Square & Hortus and New Market & China Town). Simultaneously, activity zones with high betweenness centrality are more likely to be combined with a visit to activity zones of another community (Albert Cuyp Market and Munt Square & Flower Market).

These results indicate that crowded activity zones with high betweenness centrality (e.g. Dam Square & Rokin) are in the activity chain because of the spatial proximity, while they are not necessarily important for the overall activity pattern. Important activities that structure the activity pattern have low betweenness centralities (Oosterpark and Albert Cuyp Market). These insights imply that, in general, crowdedness at activity zones with high betweenness (Dam Square & Rokin) can be alleviated by increasing the connectivity of these zones to nearby activity zones (with low betweenness, e.g. Spui) to neighbouring activity zones of Dam Square & Rokin using wayfinding signalling.

4.5.4 How are activity space compactness, travel time ratio, and activity patterns related?

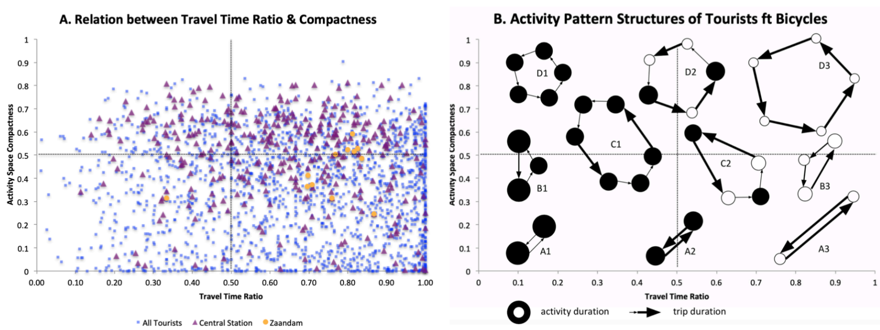
This section elaborates on the findings to reveal spatial differences in activity space. As mentioned in section **4.3.3**, the revealed activity space is based on the convex hull enclosing individual visited activity locations.

Figures 22 A displays the revealed activity spaces of all 1.817 day-tourists. The data reveals that 6.4% of the day-tourists travel to distant activity locations in the outer areas, such as Monnickendam, Zaanse Schans, Haarlem, Bloemendaal, Vinkse Plassen, and Muiden. Compared to research conducted among pedestrians, these results clearly demonstrate a far larger spatial distribution of the tourists (Shoval, Schvimer & Tamir 2018). It should be noted that the raw GPS data revealed that tourists also take the bicycle on the train to more remote cities (The Hague, Utrecht, De Keukenhof), but these trips have been excluded from this study. The widespread reach of cycling tourists offers opportunities to attract more tourists by bicycle to the outer regions of Amsterdam, such as the satellite cities Haarlem and Amstelveen, human-engineered natural parks such as the Waterleidingsduinen, and even Schiphol Airport. Notwithstanding, the majority of the tourists, however, uses the bicycle strictly within the A10 beltway and city centre of Amsterdam (pink area in Figures 22 D).



Figures 22 A-E. Activity space contours of A. all 1.817 tourists, B-C. tourists visiting the activity zone Central Station (340) and Zaandam Region (12), D 1.661 tourists staying solely within the A10 beltway, E 156 tourists bicycling to the outer areas.

Figures 21 A illustrates the corresponding relation between travel time ratio and compactness of the activity space of all tourists. The LUCY study does not indicate that increased travel time ratio relates with the compactness of the activity space of tourists. This implies that activity patterns of tourists vary from directed activity spaces, where cycling is the core activity, to compound activity spaces, where cycling is instrumental, and any combination in between (Figures 21 B). Whether this behaviour is particular for tourists, or for cyclists is out of scope for this chapter.



Figures 21 A-B. A. Relation between travel time ratio and compactness. B. Theoretic activity pattern structures based on activity space compactness and travel time ratio

We expect differences in the shape and size of the activity spaces depending on the activity zones that have been visited. In particular, differences between tourists that visited only activity zones within the city centre and tourists that (also) travelled to the outer areas of Amsterdam. Therefore, the activity spaces are also assessed for all tourists that at least visited the zones Central Station and Zaandam (See Figures 22 B-C). Tourists visiting the Central Station activity zone do usually not have a main destination, given that the compactness of their routes is often above 0.5.

In general, a high travel time ratio indicates that the bicycle trip is more important than the visit to the historic sites. The travel time ratio varies leading activity pattern structures focused on activities (C1, D2) as well as trips (B3) (as depicted in Figures 21 A-B). Tourists visiting the outer area in the Zaandam Region have more directed activity spaces, while they also divert to other activities in the neighbourhood of the main destinations if these activities are relatively close to the direct distance (C2 in Figures 21 B) between their main destinations. The high average travel time ratio thus indicates that the bicycle trip is an important activity. Thus, nice sceneries along the bicycle tracks are important for tourists.

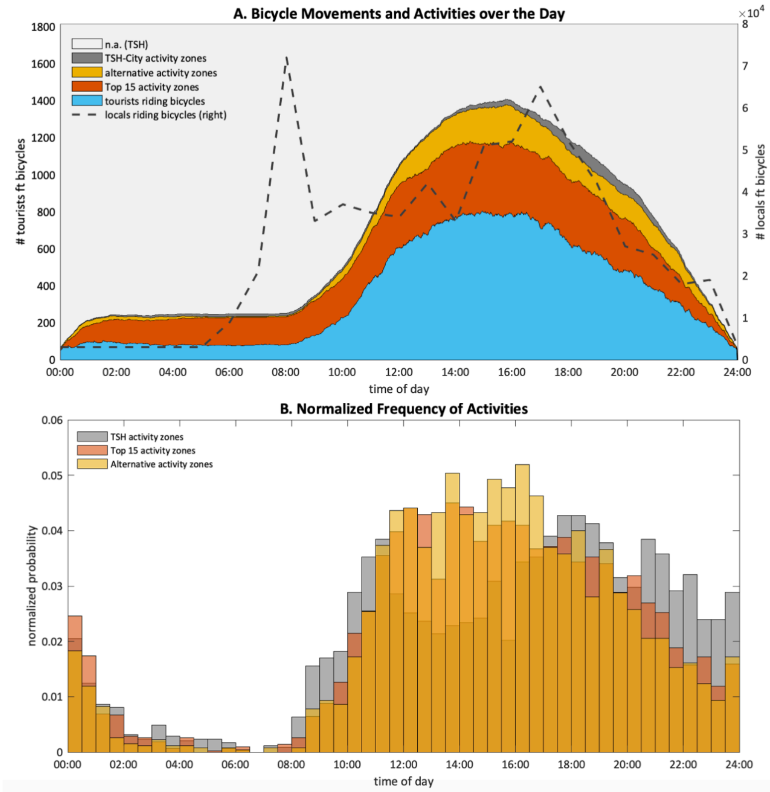
4.5.5 When do tourists cycle in Amsterdam?

Depending on the time of day, more (i.e. day time) or less (i.e. night time) tourists are expected to cycle and/or perform activities. Figures 23 A shows the travel and activity behaviour of tourists using the bicycle over time, irrespective of the day. For every minute running from 00:00 to 24:00 each tourist is riding a bicycle, visiting an activity (TSH, Top 15 activity zones, or alternative activity zones), parked before the first activity has been detected (n.a.), or parked after last activity has been detected (n.a.).

The most apparent observation is one peaked-graph with a flat tail between 00:00 and 07:00, and a sharp tail cutting of at 24:00. Gradually more tourists become active between 10:00 and 16:00. At 16:00 75% of the tourists that rented a bicycle are actively using the bicycle. After 16:00 there is a steady decrease in both activities and movements.

The Top 15 activity zones make up 74% of the activities. The graph also shows that from 13:30 to 16:30 a third of the activities are taking place outside the Top 15 most attractive activity zones. Although most of the alternative activity zones will be located within the city centre, this may indicate that, to some extent, tourists distribute themselves to other destinations when the main attractions are getting more crowded. However, behavioural insights into the decision-making process of tourists are needed to investigate why tourists would visit alternative activity zones over Top 15 activity zones.

Figure B illustrates the starting time and duration of activities of tourists over the day. The graph shows that there are differences in the activity duration depending on the type of activity (TSH, Top 15 activity zone, and alternative activity zones). The starting times of activities at TSH, exhibit 3 peaks: (i) 00:00 first detection of the bicycle with longer activity durations up to the first trip, (ii) 10:00 to 12:00 most likely the moment when most bicycles are used for the first time as the first GPS data point is always encoded as an activity, and (iii) 17:00 when most tourists return to the hotel, this is also when many museums close. After 08:00 the activity durations at TSH are decreasing. The opposite can be observed for activities performed in the city, the average activity duration is approximately 1 hour, and activities starting at 09:00 have a 75% chance to last up to 6 hours.



Figures 23 A-B. A. Bicycle movement and activities of tourists in Amsterdam over the day. Numbers of locals are based on Meerjarenplan Fiets 2017-2022 (Gemeente Amsterdam 2017). B. Starting time of activities (TSH, Top 15, and alternative) over the day.

4.6 Implications for strategies to distribute tourists flows

Recently, global policy recommendations, strategies, and measures have been established to better manage urban tourism, such as "disperse tourists within the city and beyond" (UNTWO 2018b). However, the effect of the measures heavily depends on the context and urban travel behaviour of tourists. Based on the results of this chapter several measures of the three strategies can be determined for urban bicycle tourism in the Amsterdam region. Further research is required to investigate the effectiveness of the measures. The three strategies are i) stimulate new bicycle itineraries for tourists, ii) promote spatial dispersal of tourists, and iii) promote temporal dispersion of tourists.

The first strategy advises stimulating new visitor itineraries and attractions. Ideally, this strategy aims to ensure that also local communities in outer areas benefit from tourism. To ensure local communities in outer areas also economically benefit from tourism, the possibilities and bicycle travel times should be better communicated and adapted to tourists standards. Many activity zones identified by k-means cluster analysis in the outer areas of Amsterdam are within a 45 minute bicycle trip (considering an average bicycle speed of 8 km/h) from the city centre, which is lower than the average duration of a bicycle trip of tourists that participated in the LUCY study, which is 48 minutes. Promotion of attractions and bicycle tours can be achieved through, for instance, tailored (seasonal) urban bicycle maps for tourism, including official parking places and day-itinerary suggestions to avoid the crowd, and stimulation of residents of socially deprived neighbourhoods and outer areas to organize local cycling tours during the summer period.

The second strategy suggests promoting dispersal of visitors within the city (and beyond). The activity space analysis demonstrates that 92% of the tourists prefer to remain within the city centre, while there is also the potential to distribute tourists to outer areas by bicycle. Pop-up events and new markets located within ~2,23 kilometres (i.e. the average Euclidean distance between activities) from an activity zone of the Top 15 are potential attractions to distribute tourists within the city while travelling by bicycle. Another measure that has been proposed by UNTWO is to establish a strong joint identity. However, the analysis of activity sequences revealed that tourists also seek variety. Therefore, it is important to create agglomerations of activity zones with similar urban identities to allow them to be distinct as a salient region. At the same time, it is important to ensure that a certain degree of variation exists within the salient region. Additionally, within a Top 15 activity zone visitors can be distributed to alternative attractions at walking distance (e.g. near Museum Square activity zone there are the less famous Amsterdam Art Station, Zuiderbad, Café Loetje and Wildschut). This measure will relieve the crowdedness at main attractions and increase the activity duration within activity zones, as visitors might visit both the main attraction and the secondary attraction.

The third strategy aims to promote the time-based dispersal of visitors. The results from the time patterns of tourists revealed a strong decrease in activity and movement intensities after 17:00. It is important to avoid big tourist flows during the morning and evening peak of residents. More of the major museums could explore longer opening hours (10am to 10pm), potentially even for a reduced price if tourists are travelling by bicycle. This could stimulate sustainable, healthy, and inclusive activity patterns, where tourists can visit more activity zones as alternative activity zones in outer areas can be visited at the start of the day. Further research is needed to assess the experience gained when exploring a city by bicycle and related benefits for residents as well as tourists.

The LUCY project can also be used to study movement patterns of tourists in greater detail to improve urban infrastructure and facilities. Based on the insights of activity patterns, the municipality can be advised to explore three measures. First, ensure that major routes between connected activity zones are well equipped for bicycle traffic of tourists and residents (slow speed/recreational paths and high speed/efficient paths), and that good wayfinding systems are in place. Third, capacity issues concerning bicycle-parking places can be evaluated based on the identified activity zones while incorporating the expected growth of both commuters and tourist volumes.

4.7 Conclusion

This chapter aims to increase the understanding of tourists' urban travel behaviour when travelling by bicycle leveraging on a large-scale empirical GPS data collection in Amsterdam, The Netherlands. The results are used to identify how new itineraries (activity sequences) can be stimulated, and how tourists can be better distributed spatially and temporarily. These insights are relevant for policymakers and urban planners to design, test, and evaluate their policies, design, and travel information to realize and maintain "bikeable" cities for citizens, as well as, tourists.

The research objective is to identify spatial and temporal travel and activity patterns of tourists with access to a bicycle in a metropolitan area and to see which metrics can be used to characterise these patterns. Data processing approach that was used in this chapter classified stationary and moving GPS points that have been collected during July 1st and September 1st, yielding 10.342 activity locations and 8.525 trips, made by 1.817 unique tourist days. This information has been used in four analyses to unravel spatial and temporal travel and activity patterns of tourists: k-means clustering identified 105 activity zones, network analysis identified spatial relations between activity zones based on four communities and betweenness centrality of the Top 15 most visited activity zones, activity space analysis investigated the spatial dispersal of tourists, and temporal profiles identified which moments of the day most tourists are visiting activities or bicycling in and around Amsterdam.

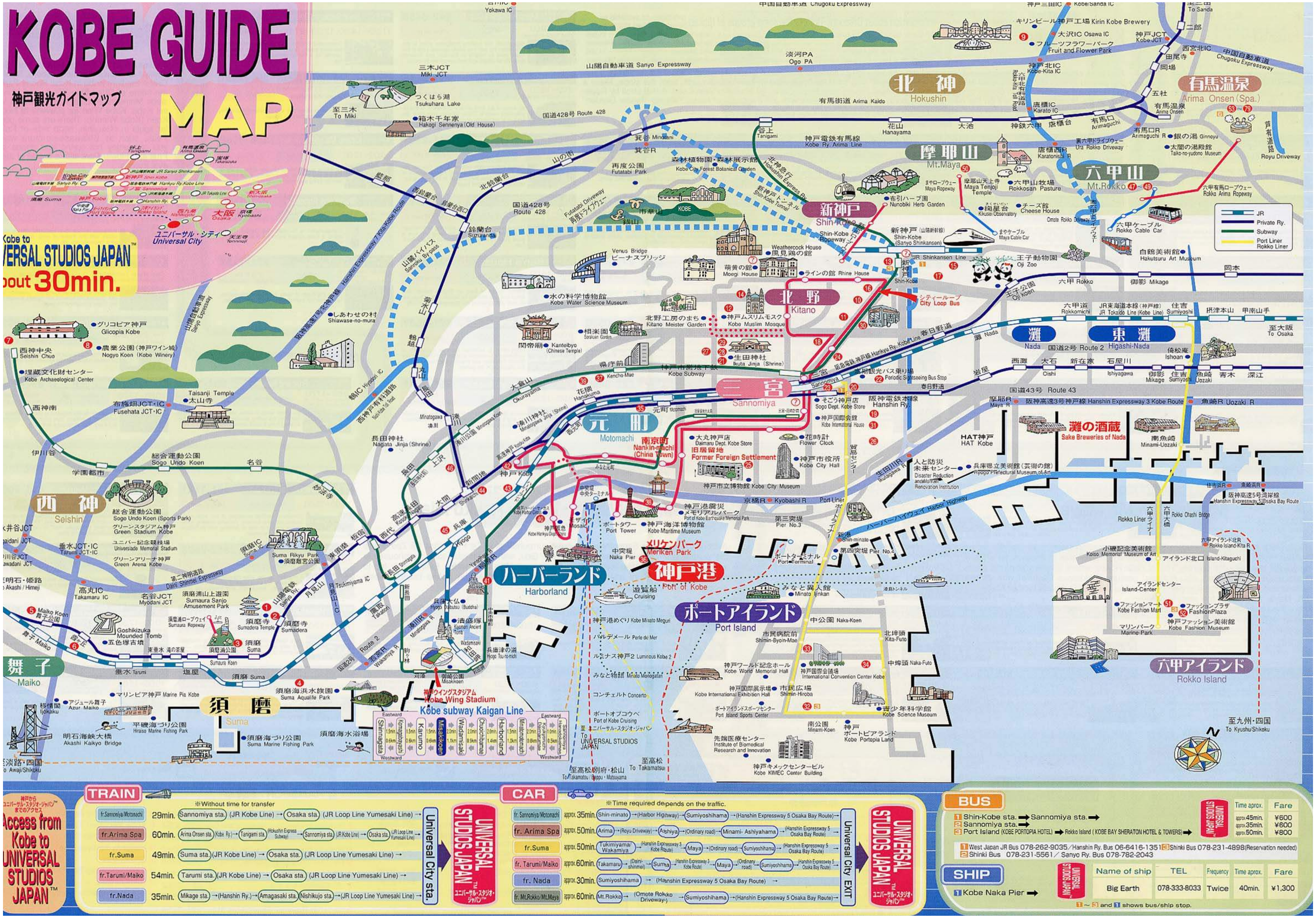
The main contribution for practice is that it is possible to achieve wider spatial dispersal of tourists if they travel by bicycle, provided that activities in outer areas are promoted at hotels, chapter and online tourists maps are available with the bicycle infrastructure and a clear overview which activities are feasible to visit by bicycle during one, two, or three days. The results also suggest wayfinding systems at A-locations indicating bicycle times for tourists, if they are lower than the expectation of tourists. If the outer areas belong to different (same) activity communities, the locations with high betweenness centrality (clustering coefficient) have the most chance to disperse tourists to other communities. To alleviate local crowdedness, activity locations within a radius of 2-3 kilometres from the most crowded sites, are potential spill-over zones. Secondly, the location of hotels appears to influence the activity pattern and travel behaviour of tourists travelling by bicycle. Stimulation to built new hotels closer to the A10 could increase the visits at less common activities within the 2-3 kilometre radius.

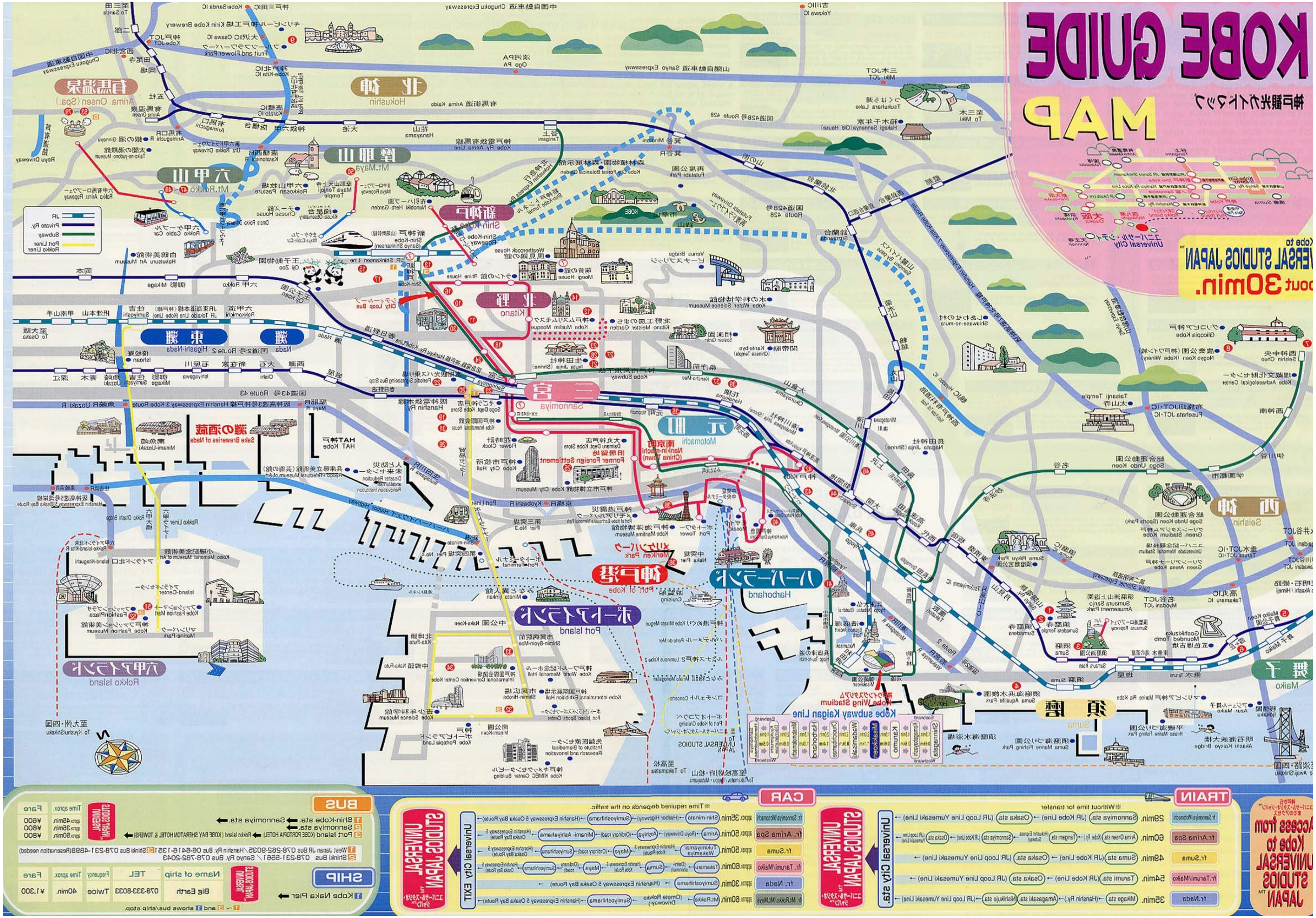
The main scientific contributions are the insight that the activity network of tourists with access to a bicycle is consistent with expectations from other transport networks, there are differences between communities, and correlations with betweenness centrality are positive for closeness centrality and weighted degree, while negative for the clustering coefficient. Secondly, the combination of activity spaces, travel ratios, and travel pattern structure extend existing theories of tourists' travel behaviour (e.g. McKercher & Lau 2008). Mobility patterns of tourists using bicycles in metropolitan regions vary from activity oriented to trip orientated, and from directed to compact. Further research is required to explore the difference between citizens and tourists, and tourists travelling by foot and public transport (Chapter 5).

A limitation of this chapter relates to the large-scale GPS data collection of bicycle movements; for privacy reasons, no personal data has been collected. Therefore, this chapter does not provide insights into the tourist and/or activity typologies that can be used to tailor measures to specific user groups. Another limitation of the current LUCY study is the strong bias towards the start and end of all day trips at TSH. To make more generic theories and statements, more origins (hotels) should be included in the data collection process. If future research could combine both GPS travel itineraries and complementary surveys and experiments, more promising research questions can be answered such as 1.) How do tourists learn the structure of the spatial environment they are travelling through, 2.) How do they memorize what they have learned over time, and 3.) How do they utilize memorized knowledge to find the way in large-scale urban environments?

Acknowledgements

This research was supported by the Allegro project (no. 669792), which is financed by the European Research 8 Council and Amsterdam Institute for Advanced Metropolitan Solutions.


KOBE GUIDE MAP
神戸観光ガイドマップ
Kobe to UNIVERSAL STUDIOS JAPAN about 30min.
北神 Hokushin
摩耶山 Mt.Maya
六甲山 Mt.Rokko
有馬温泉 Arima Onsen (Spa.)
新神戸 Shin-Kobe
北野 Kitano
三宮 Sannomiya
元町 Motomachi
南京町 Nankin-machi (China Town)
旧居留地 Former Foreign Settlement
灘 Nada
東灘 Higashi-Nada
灘の酒蔵 Sake Breweries of Nada
西神 Seishin
須磨 Suma
舞子 Maiko
ハーバーランド Harborland
神戸港 Port of Kobe
メリケンパーク Meriken Park
ポートアイランド Port Island
六甲アイランド Rokko Island
Kobe subway Kaigan Line
神戸ウイングスタジアム Kobe Wing Stadium
JR
Private Ry.
Subway
Port Liner Rokko Liner
TRAIN
CAR
BUS
SHIP
Access from Kobe to UNIVERSAL STUDIOS JAPAN
Universal City sta.
Universal City EXIT
fr.Sannomiya/Motomachi 29min. Sannomiya sta. (JR Kobe Line) → Osaka sta. (JR Loop Line Yumesaki Line) →
fr.Arima Spa 60min.
fr.Suma 49min. Suma sta. (JR Kobe Line) → Osaka sta. (JR Loop Line Yumesaki Line) →
fr.Tarumi/Maiko 54min. Tarumi sta. (JR Kobe Line) → Osaka sta. (JR Loop Line Yumesaki Line) →
fr.Nada 35min. Mikage sta. → (Hanshin Ry.) → Amagasaki sta. → Nishikujo sta. → (JR Loop Line Yumesaki Line) →
※Without time for transfer
※Time required depends on the traffic.
1 Shin-Kobe sta. → Sannomiya sta. →
2 Sannomiya sta. →
3 Port Island (KOBE PORTOPIA HOTEL) → Rokko Island (KOBE BAY SHERATON HOTEL & TOWERS) →
Time aprox. approx.45min. approx.35min. approx.50min.
Fare ¥600 ¥600 ¥800
1 West Japan JR Bus 078-262-9035／Hanshin Ry. Bus 06-6416-1351 3 Shinki Bus 078-231-4898(Reservation needed)
2 Shinki Bus 078-231-5561／Sanyo Ry. Bus 078-782-2043
1 Kobe Naka Pier →
Name of ship: Big Earth | TEL: 078-333-8033 | Frequency: Twice | Time aprox.: 40min. | Fare: ¥1,300
1～3 and 1 shows bus/ship stop.



KOBE GUIDE
MAP
神戸観光ガイドマップ
Kobe to UNIVERSAL STUDIOS JAPAN about 30min.
Access from Kobe to UNIVERSAL STUDIOS JAPAN
TRAIN
CAR
BUS
SHIP
Universal City sta.
Universal City EXIT

5

On the Relation between Learning the City and Routing

To support policy makers to gain more insights into the impacts of urban tourist flows, a greater understanding is required in the detailed urban travel behaviour. Based on the same GPS data as presented in chapter 4, this chapter presents the dynamics of the spatial behaviour when new knowledge is acquired with every bicycle trip of tourists. This is done by modelling the influence of experienced travel behaviour and spatial knowledge acquisition on the "route selection space" of tourists travelling by bicycle in Amsterdam. Generalized Estimating Equations (GEE) are used to assess the (spatial) learning effect of tourists as a function of the approximated trip purpose, familiarity, and movement patterns. Four route selection space characteristics are investigated in this contribution: detour ratio, maximum deviation from the bearing line, eccentricity, and curvature. The five highlights are:

1. Investigate route patterns of city tourists travelling by bicycle.
2. Route selection space dynamics analysed by spatial probability distributions and Generalized Estimating Equations.
3. Route selection space of tourists mainly depends on trip purpose.
4. Tourists learn within a day through the number of trips and new activities.
5. Provide behavioural insights to improve management of urban tourism.

This is an edited version of the following article:
Zomer, Duives, Cats, and Hoogendoorn (under review). On the Relation between Learning the City and Routing. *Transportation Research Part F: Travel behaviour and Psychology.*

5.1 Introduction

The Netherlands can expect a growth of 44%-200% in the number of tourists, yielding 28,8 to 41,9 million tourists in 2030. Currently there is an unequal dispersion of tourists over the country, as 40% (8 million) of the tourists stay in the capital city, Amsterdam. If the same share of tourists remain in Amsterdam in 2030, the city centre could be facing "overtourism" as for decades the main destinations for tourists are strongly concentrated in a triangle between the Central Station, Vondelpark and Weesperbuurt (e.g. Jewisch Quarter & Hermitage Museum). Overtourism creates tension between citizens and tourists that decreases the quality of life of both due to excessive noise, nuisance, and pressure on infrastructure (UNTWO 2018). The effects of global strategies and measures to better understand and manage urban tourism heavily depends on the travel behavior of tourists within the respective cities.

The bicycle is the main mode of transportation in Amsterdam and it seems from the street view and media attention it is getting more popular amongst tourists. The bicycle offers sustainable, healthy, and cheap opportunities to disperse tourists to outer areas and alternative destinations within the city (Chapter 4). However, it is hypothesized that there are many differences between the routing behaviour of tourists and commuters: commuting behaviour is characterized by i.) habitual choices with limited variety (Lima et al. 2016), ii.) the fact that spatial knowledge is matured as new routes and destinations are rare, and iii.) travel is seen as a means to perform a certain activity at the destination with travel time ratios of 0,1 (Dijst & Vidacovic (2000). As the travel time ratio is derived by dividing the median travel by the sum of the median travel time and median activity duration, this means that commuters are willing to travel 1 minute for an activity duration of 10 minutes. The travel time ratio of 0,6 reveals that tourists are willing to travel for 20 minutes by bicycle for a 10 minute activity in Amsterdam (Chapter 4). This difference implies that for tourists the bicycle trip might be more important than the activity, while commuters consider travelling merely a necessity to perform the more important activity.

We hypothesize that tourists first perform explorative wayfinding to compose a choice set and some basic understanding of the city they are travelling in. With every new trip they may explore more, acquire more spatial knowledge and/or retrace their steps. The trip purpose is on average more important than the activity with a travel time ratio of 0,6 (Chapter 4). The latter means that tourists are willing to travel for 20 minutes by bicycle for 10 minutes of activity. Due to these differences it may be prohibitive to use the same activity and route choice models to model and predict bicycle behaviour of tourists. To advance the understanding of bicycling behavior of tourists, insights into activity and movement patterns of tourists and how choices and patterns evolve over time are needed. Therefore we need to model the influence of experienced travel behaviour and spatial knowledge acquisition on the route selection space of tourists travelling by bicycle in Amsterdam.

Spatial behavior of tourists can be described as a direct function of their experience with the built environment (Golledge & Stimson 1987). To the best of the authors' knowledge, the development of spatial behavior when travelling in an unfamiliar environment is largely unknown, especially when these movements are performed by bike. The aim of this chapter is to gather insights into the impact of experience with the built environment on tourists' spatial behaviour. Our research question is thus formulated as follows "To what extent does experience with the built environment influence the development of spatial behavior of tourists by bicycle over time?". Here, spatial behaviour depends on direct distance between origin and destination and trip purpose characteristics. While tourists' experience is quantified by means of the trip number, their historic travel experience, and previously acquired spatial knowledge and routing behaviour.

Spatial choice sets are the result of a complex interplay between spatial restrictions, activity space, and personal abilities and preferences (Bovy & Stern 1990; Manaugh & El-Geneidy 2012). One of these spatial choice sets, namely the *route selection space* is determined in this chapter using (the development of) detour ratios, maximum deviation, eccentricity, and curvature (over time). These four metrics are used to describe the efficiency of movement patterns and can jointly be used to define a boundary on the set of feasible route alternatives when generating a choice set. The main contribution of this chapter is expanding our understanding of activity and route patterns in relation to the development and impact of (spatial) familiarity of tourists on the four metrics. Moreover, our study provides insights for policy makers that can be used to improve tourist activity and routing information aimed to attract tourists that have access to bicycle to less crowded areas.

The outline of this chapter is as follows. In the next section we provide a literature review on the route selection space and knowledge acquisition (Section 5.2). In Section 5.3 the data collection methods are discussed. This is followed by an elaboration on the research approach in Section 5.4. Section 5.5 presents insights into the dynamics of the route selection space and summarizes the descriptives of the dependent and explanatory variables. Four models are discussed in Section 5.6 to describe the dynamics of the route selection space. The implications of the findings from Section 5.5 and 5.6 are concluded and discussed in Section 5.7.

5.2 Literature on route selection space & knowledge acquistion

This section discusses the state-of-the-art relating to spatial choice sets and knowledge development in section 5.2.1 and section 5.2.2, respectively. The section concludes with the expectation of the relations as expected from the existing theories in section 5.2.3.

5.2.1 Spatial choice sets

The origin of the spatial route choice set concept can be found in Hägerstrand's space-time geography (Hägerstrand 1953). The *Potential Path Area* (PPA) is the projected ellipse of the space-time diagram on the surface, which represents all locations that a person can occupy during the time available between two sequential activities (t_i, t_{i+1}) (Miller 2005). The difference is that the PPA represents spatial behaviour at trip level, while the spatial route choice set represents spatial behaviour of individuals. Another notion, more common in literature and closely linked to the spatial route choice set, is the *activity space*.

Similar approaches have been used to represent individual and household activity spaces, for instance using ellipses (Newsome, Walcott & Smith 1998), minimum spanning trees and kernel densities (Schönfelder & Axhausen 2002) and local travel index (Manaugh & El-Geneidy 2012). Model results and significant determinants reported in these four studies are documented in Table 12. Schönfelder & Axhausen (2002) reflect on these methods, and conclude that activity space ellipse overgeneralizes the spatial pattern leading to an oversized area, kernel densities ignore connections between activity locations, and minimum spanning tree only captures the spatial distribution of the activities. They propose to combine the minimum spanning tree with a spatial buffer to incorporate the size of human activity spaces, called the road network buffer approach.

Table 12. Detailed findings of model results of choice set ellipse in literature.

Details	Dependent variables					
	deviation ratio	*ellipse area*	*total elapsed time*	*minimum spanning tree*	*kernel density*	*Local Travel Index*
ANOVA, 653 chained work trips. Charlotte (USA)	Urban typology Income Household size	Not significant	Urban typology Age Household size Race			
GLM, 126/132 respondents reporting activity travel behaviour for 6 weeks, conducted more than 40 trips. Halle & Karlsruhe (Germany)				Distance home-city centre Age Main car user PT subscription # unique activities # trips	Urban typology Distance home-city centre Age Main car user PT subscription Income # unique activities # trips	
Regression analysis, 11.633 households. Montréal (Canada)						Regional &local accessibility Household type # trips Trip purpose % walk trips
Probability density function, GPS routing data of 526 private cars over an 18-month period		Detour that people are willing to take is bounded by an elliptical shape Human routes have lower eccentricity compared to optimal routes Eccentricity does not depend on trip distance				

[1] Newsome, Walcott & Smith (1998), [2] Schönfelder & Axhausen (2002), [3] Manaugh & El-Geneidy (2012), [4] Lima et al. (2016).
Variable categories: blue built environment, red individual or household, orange travel behaviour

Only one study has analyzed the existence of the route selection space (RSS) based on a large data set of car drivers, which was coined the boundary of human routes (Lima et al. 2016). Controlling for direct distance and direction, they found that 95% of all routes by car do not deviate more than direct distance left or right from the bearing and half the direct distance backwards and beyond the origin and destination. Most of the routes are contained in an ellipse around two foci points (origin and destination), which implies that most detours are well-bounded. To describe the dispersion from the bearing line, they use a measure called eccentricity, which is defined by the deviation between direct distance and the maximum value of the sum of the geodesic[1] distance between the origin and destination (Lima et al. 2016). They compared their findings with eccentricities from optimal routes and concluded that human routes have wider spatial route choice sets. Furthermore, they found indications that the RSS of car trips is independent of the Euclidean trip distance.

Next to the network layout, also other variables have been identified to impact the RSS. For instance, Bovy & Stern (1990) hypothesize that subjective spatial restrictions, personal preferences, and activity patterns determine the boundary of the RSS, leading to individual route selection spaces, while Golledge & Stimson (1987) developed a theory that demonstrates that spatial behavior is a direct function of their individual experience with the built environment. Yet, the dynamics of the individual route selection space when familiarity is under development and the relation with travel behaviour remain known.

5.2.2 *Spatial knowledge development*

A second element that is important in finding the way in large-scale spaces is environmental cognition. That is, space must be cognitively organized and memorized when the entire route cannot be perceived at once, or when all feasible routes cannot be perceived as a sequence of discrete views (Stea & Blaut 1973). Environmental cognition consists of spatial knowledge of locations (distance, direction, and relative relation) and associated (descriptive and evaluative) attributes. The latter, namely associated attributes, are dependent on the measurement scale (e.g. country, city, shopping mall). There is a long-lasting hypothesis without consistent evidence that assumes knowledge of individuals' cognitive maps can be used to predict their spatial behaviour (Fishbein 1967).

This human knowledge of cognitive maps consists of perceptions and memories and can deviate from reality leading to spatial cognition distortions. A meta-literature review, covering results from various experiments, conducted almost three decades ago identified differences in relative accuracy of cognitive distance (Wiest & Bell 1985). Immediate distance observations of respondents are on average 8% higher, while memorized previously visited destinations and inferred distances to unknown destinations are 9% and 25% smaller compared to the actual direct distance (Wiest & Bell 1985). There are also many inconsistencies in the literature; some studies revealed that city centre destinations are perceived smaller compared to remote destinations (at equal direct distance from the observer) (Lee 1962; 1970), while other studies found that innercity trips are perceived relatively longer compared to longer remote trips (Allen 1981; Montello 1997).

Familiarity can evolve with every trip and activity and affects activity patterns and route choices. However, spatial and network knowledge is only acquired when experiences of previous trips and activities are processed and memorized. Moreover, the perception of attributes becomes more accurate when the acquired knowledge is appropriately applied to future and new activity and route choices (Stern & Leiser 1988). The ability to process and apply the newly acquired knowledge (directly) to future trips depends also on individual spatial abilities and preferences. Diminution and memory loss or selection ensures that excess information is lost and important features are retained (Miller 1956). Limited memory retention has been modelled in a cognitive learning model of daily activity-travel patterns based on the shortest path and attention and sensitivity to environmental attributes (Cenani, Arentze & Timmermans 2012).

[1] Geodesic distance is our study defined as the shortest path consisting of two connected edges between origin, one of the points along the observed trajectory, and destination.

5.2.3 *Expectations of development of tourists' route selection space*

In literature we find various constructs related to the route selection space, and also explanatory variables (i.e. urban typology and accessibility, individual or household characteristics, and number of trips, trip purpose or unique activities) have been identified. There is also a coherent theory on development of spatial familiarity, yet the dynamics of the individual route selection space when familiarity is under development and the relation with travel behaviour remain known.

Tourists start without spatial or network knowledge of the built environment, but according to the accretion principle their familiarity develops already after the first trip (Stea & Blaut 1973). To find the way to the next activity location the acquired knowledge can influence the detour ratio and deviation area of the next trip. This depends on trip length, size of new and old areas that have to be explored and retraced, and time pressure.

5.3 Data collection

To derive tourists' activity and movement patterns two months of GPS data from 250 bicycles of The Student Hotel Amsterdam-City (TSH) has been used for the months of July and August of 2017 (Chapter 4). During this period TSH houses only Hotel Guests, which are predominantly tourists visiting Amsterdam. These tourists have the possibility to rent a bicycle during their stay and were given the option to voluntarily participate in the LUCY study. At the beginning of their bicycle rental period, it is assumed that all tourists have only experience of The Student Hotel local surroundings upon embarking on their first bicycle trip in Amsterdam.

When using the bicycle, every 10-30 seconds the longitude, latitude, speed, and a timestamp are recorded. Between July 1st 06:00:00 and September 1st 06:00:00 a total of 1.465.590 GPS points have been collected from 250 bicycles equipped with the trackers. Filtering and processing procedures have been applied to the data to reduce inaccuracies and enrich the raw data. The raw GPS data require substantial cleaning efforts due to fluctuations in time and space caused by low battery power, characteristics of the built environment (e.g. high rise with glass facades), very low travel speeds, and historic cache memory of the tracker. As part of this process, heuristics were developed to identify each data point as part of an activity, a movement or an outlier. For details, we refer to (Chapter 4).

The result of the filtering and processing procedures is a set of 1.810 unique tourists with 8490 trips. The average trip takes 20 minutes, have a Euclidean distance of 1,64 km, a travel distance of 8,5 km, and takes place between 11am in the morning and 7pm in the evening.
Tourists mostly move around and inside the city centre of Amsterdam and occasionally travel to the surroundings of the city of Amsterdam, including some regional villages, such as Zaanse Schans and Bloemendaal aan Zee.

5.4 Research approach

The research approach consists of two parts: representing the spatial probability of the route selection space, and the identification of learning effects on four characteristics that approximate the route selection space.

Below, the research methodology to determine the impact of spatial knowledge on the route selection space is further elaborated upon. Before doing so, an representation of the route selection space is presented (5.4.1), illuminating the relation between the ellipse that bounds the route selection space and four metrics, namely the detour ratio, the maximum deviation, the eccentricity, and the curvature. In section 5.4.2 we explain the derivation of the explanatory variables that describe the dynamics in movement patterns of tourists. The final section discusses the approach to model the influence of spatial learning on routing patterns.

5.4.1 *Quantifying the route selection space*

The spatial route choice set of an individual tourist can not be observed directly, also referred to as latent. Therefore, route selection space characteristics are needed to approximate its shape and dynamics. In this chapter four characteristics are investigated, namely the detour ratio, the deviation area, the maximum deviation from bearing line, and the eccentricity (Figure 24). A comparison between trips is only possible when the trajectories of all trips are normalized with respect to scale and direction. Findings from the study of Lima et al. (2016) indicated that the shape and dimensions of the route selection space of commuters is independent from the distance between origin and destination. Therefore, the Euclidean trajectory coordinates for each *trip*$_i$ of *tourist*$_n$ are transformed to Cartesian coordinates, with all origins at location (0,0) and destinations at location (0,1). Afterwards, the four characteristics are computed using the geodesic (normalized) distances.

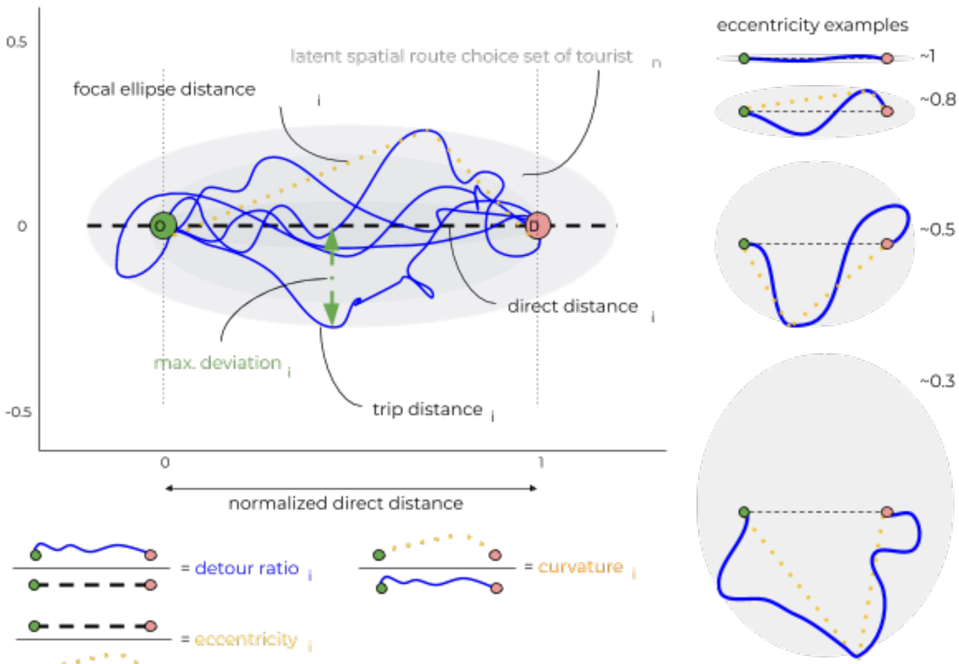


Figure 24. Illustration of the foure route selection space characteristics (by authors).

5.4.1.1 *Detour ratio*

The detour ratio (route factor) describes the relation between the actual travel distance of a trip and the direct distance between the origin and destination. The detour ratio of tourist *n* and trip *i* is defined as:

$$DetourRatio_{ni} = \frac{trip\ distance_{ni}}{direct\ distance_{ni}} \quad (7)$$

The minimum value is 1, corresponding to a travel distance equal to the direct distance.

5.4.1.2 *Maximum spatial deviation*

The maximum deviation from the bearing line in the Y-axis describes how far a tourist is willing to deviate, regardless of the position on the X-axis. This metric is computed to find the maximum euclidean distance between thenormalized Y coordinates of all interval points along the trajectory of trip i.

$$MaxDevation_{ni} = \max(Y_{ni}) \tag{8}$$

5.4.1.3 *Eccentricity*

The eccentricity is similar to the metric used in a study on individual routing behaviour of car drivers (Lima et al. 2016). It measures the shape of the route selection space based on the direct distance and the focal ellipse distance (the maximum Euclidean distance that is needed to reach any point on the trajectory from the origin and destination), see eq. 9 for the mathematical description of this metric. The focal distance represents the minimum trip distance that is required to have a similar route selection space size. The examples in Figure 24 show that a value of 1 corresponds to a trip distance equal to the direct distance, while a value of 0.5 corresponds to a circular shape and even lower values represent an elliptical shape where the width is larger than the direct distance between origin and destination.

$$Eccentricity_{ni} = \frac{direct\ distance_{ni}}{max\left(distance_{OF_{ni}} + distance_{DF_{ni}}\right)} \tag{9}$$

5.4.1.4 *Curvature*

The curvature is an original metric that we hereby introduce. It is based on the detour ratio and the eccentricity. Curvature describes the relation between the focal distance and trip distance, with a minimum value of 0 and a maximum value of 1. A value of 1 represents the most efficient route possible, considering the size of the route selection space, while lower values indicate that there are more deviations from the focal distance axes.

$$Curvature_{ni} = \frac{max\left(distance_{OF_{ni}} + distance_{DF_{ni}}\right)}{trip\ distance_{ni}} \tag{10}$$

5.4.2 *Analyzing evolution of dynamics in tourists movement patterns*

In order to analyse the development of tourist movement patterns, one has to quantify the movement patterns. This section defines the variables that will be used in the subsequent analysis to quantify the historic, current and future movement patterns of cyclists. Here, it is important to note that this chapter assumes that all previous actions add to the current spatial knowledge of the individual. Thus, this chapter does not account for memory decay.

First, section 5.4.2.1. details the characteristics of the next trip and destination. Subsequently, section 5.4.2.2 presents the variables that describe all previous trips and performed activities.

5.4.2.1 *Next trip and destination*

The Euclidean direct distance of $Trip_i$ is used to represent the spatial dispersion from the current location, the origin. The travel time ratio is derived by dividing the travel time by the sum of the travel time and activity duration of the coming trip (Dijst & Vidakovic 2000).

Third, a counter monitors if the destination is located in previously visited area coined as covered area (see eq. 11).

5.4.2.2 Memorized trips and activities

Familiarity with the environment can be approximated by a function of the total time spent in the city (exposure time), and depends on the memorized cognition of perception and experience. The familiarity is hierarchical and based on previous activities and routes. In this chapter familiarity is quantified by a combination of several attributes, namely the travelled distance & time, the covered area, the retraced area, the space variation index and the number of retraced regions. Underneath, the definition of each of these attributes is detailed.

5.4.2.2.1 Travelled distance, travelled time and median travel speed

The travelled time and distance represent the time spent and distance covered from the moment in time tourists start cycling to the current moment in time. These two attributes are the cumulative summation of the distances and time covered during all previous trips. There is no memory of travelled distance and time when modelling the first trip, hence the initial values are set to 0. Activity duration is the sum of all previous activity durations. Upon embarking on the first trip only the activity duration of the respective activity is taken into account. Memorized travel speed is taken as the median travel speed over all previous trips. Since there is no memorized travel speed before the first trip, the memorized travel speed for the first trip is set to 0.

5.4.2.2.2 Covered area, retraced area and variation index

To compute the covered, retraced, and experienced area, with each new trip a new polygon is created that consists of the union of a 100 metre buffer around $Trip_{i-1}$ and a 300 metre buffer around $Activity_i$, where the intersection between $Trip_{i-1}$ and a 300 metre buffer around $Activity_{i-1}$ is excluded: $\left(\frac{BufferedTrip_{i-1}}{BufferedActivity_{i-1}} \cap BufferedActivity_i\right)$.

Another set of attributes is added to describe the covered area, the experienced area, the retraced area, and the space variation index. In doing so, a structure of multiple layers of multiple buffers is used to compute these four attributes coined as RetraceAreaLevel$\{l\}$. RetraceAreaLevel{1} for $Trip_i$ consists of the union of all previously created polygons, and as such, contains the areas covered at least once, while RetraceAreaLevel{2} contains the areas covered at least twice, and RetraceAreaLevel{3} contains the areas covered at least three times. As such, RetraceAreaLevel$\{l\}$ is visited at least l times. The RetraceAreaLevel structure has the property to only add another level if a new trip intersects with an area that is currently the maximum level l. Here, the covered area consists of the unique area that has been visited upon embarking on the next trip.

$$CoveredArea_{ni} = \text{area}(RetraceAreaLevels_{ni}\{1\}) \qquad (11)$$

Additionally, a counter monitors if the destination is located in previously visited area (covered space), and how many times.

The experienced area consists of the summation of all areas that have been visited upon embarking on the next trip. At the first trip both the covered area and experienced area have the size of the buffer space around the first activity $\left(\frac{300^2\pi}{4}\right)$.

$$ExperiencedArea_{ni} = \sum_{1:l} \text{area}(RetraceAreaLevels_{ni}\{max(l)\}) \qquad (12)$$

The retraced area is defined as the difference between the experienced area and the covered area.

$$RetracedArea = ExperiencedArea - CoveredArea \tag{13}$$

The space variation index details the relation between the retraced area and the covered area. Therefore, it provides an indication about the tendency to explore or retrace. A value of 0 implies that the tourist has only been exploring new area during the previous trips and activities, i.e. choice set formation phase. A value of 1 implies that the tourist always retraces previous routes and activities, i.e. the spatial knowledge acquisition is mature.

$$SpaceVariationIndex = \frac{RetracedArea}{CoveredArea} \tag{14}$$

The number of anchor areas describes the fragmentation of the retraced area, as we assume that retraced areas have a bigger influence on the (structure of) memorized space. The number of anchor areas are computed as the number of separate regions in the area that is retraced at least once (RetraceAreaLevels$_{ni}$$\{2\}$).

$$countAnchorAreas = RetracedAreaLevels\{2\}.NumRegions \tag{15}$$

5.4.3 *Modelling route selection space metrics*

The development of familiarity is hypothesized to influence the detour ratio, maximum deviation, eccentricity, and curvature of the next trip. This chapter investigates the impact of different conceptualizations of familiarity using the above mentioned attributes. *Generalized Estimating Equation* (GEE) models are used to assess if the movement patterns of tourists become more efficient when the familiarity with the built environment grows, which leads to a decline in detour ratio and maximum deviation and increase of eccentricity and efficiency of the curvature.

5.4.3.1 *Generalized Estimating Equations*

Generalized Estimating Equation (GEE) are developed to analyse longitudinal and/or correlated data (Liang and Zeger 1986). This approach is conceptually different from multilevel and hierarchical models as GEEs do not explicitly model the variation. Instead it estimates the similarity of the observations (Hanley et al. 2003; Ballinger 2005). As a result GEEs are marginal models, they model a population average. The results should be interpreted as with every unit increase of an explanatory variable across the population corresponds to the change in the average response of the dependent variables. GEE models are based on (Liang and Zeger 1986):

1. A marginal model:

$$\mu_{ni} = E(y_{ni}) \tag{16}$$

2. The linear predictor:

$$\eta_{ni} = X_{ni}^{i}\beta \tag{17}$$

Where X_{ni} is the covariate vector for tourist n at trip i.

- The systematic component used to relate the response variable to the linear combination of the covariates with an identity link function.

$$g(\mu_{ni}) = \eta_{ni} \tag{18}$$

- The random component is captured by the variance as a function of the mean, and consequently the distribution of the response variable,

$$Var(Y_{ni}) = \phi V(\mu_{ni}) \tag{19}$$

- The autoregressive 1 (AR1) working correlation structure of the clustered and longitudinal response variables differentiates the GEE from a GLM.

5.4.3.2 Model structure

The efficiency of the spatial bicycle behaviour of tourists can be measured per trip using characteristics that describe the route choice ellipse. In this chapter four characteristics are investigated: detour ratio, maximum deviation from bearing, eccentricity, and curvature. The efficiency of a trip depends on the Euclidean (direct) distance and the developed familiarity. Familiarity with the environment can be described as a function of total time spent in the city (exposure time), and depends also on the travelled speed, distance, perception and experience. In our study the development of familiarity is hierarchical and based on previous activities and routes:

$$\eta(RouteSelectionSpaceIndicator)_{ni} = \beta_1 Trip_{EucDis_{ni}} + Familiarity_{ni} \qquad (20)$$

Where Familiarity (*F*) depends on the characteristics of the next trip, travele xperience from previous trips and activities, acquired spatial knowledge from previous trips and activities, and movement pattern during the previous trips:

$$\begin{aligned} F_{ni} = {} & \beta_2 Trip_{new|recDestination_{ni}} + \beta_3 Trip_{nTTR_{ni}} + \beta_4 Trip_{Sequence_{ni}} + \cdots \qquad (21) \\ & \beta_5 Exp_{bicycleDist_{ni}} + \beta_6 Exp_{bicycleDur_{ni}} + \beta_7 Exp_{activityDur_{ni}} + \cdots \\ & \beta_8 Know_{coveredS_{ni}} + \beta_9 Know_{retracedS_{ni}} + \beta_{10} Know_{\#newAct_{ni}} + \cdots \\ & \beta_{11} Know_{\#recAct_{ni}} + \beta_{12} Know_{\#anchorAreas_{ni}} + \cdots \\ & \beta_{13} Move_{medSpeed_{ni}} + \beta_{14} Move_{spaceVariation_{ni}} \end{aligned}$$

Generalized Estimating Equations are used to identify which determinants of familiarity development influence the four measurements of efficiency.

5.5 Descriptive results

Before the model results are discussed in Section 5.6, first the descriptive results are explained. The behavioural theory on (spatial) familiarity as described in 5.4.2.2 is based on previous activities and routes that have been visited. In order to understand the learning effects of tourists, we address thefollowing question: *To what extent is spatial familiarity important to predict the four spatial characteristics of the route selection space of the next bicycle trip of tourists, when the origin and destination are known, and hence the direct distance?* As mentioned above, the dataset consists of 1.810 day-tourists, who made 8.490 bicycle trips in Amsterdam. Most tourists visited 2-3 activities on a single day, and cycled 20 minutes and travelled 8,5 km per trip. Data analysis shows that the average detour ratio has a high variability for shorter trips (< 3 km), while the detour ratio to more distant activities remains constant around 1,5. The median observed time of a tourist is 3 hours and 56 minutes. Current findings also show high variation between activity spaces and bicycle movement patterns.

Behavioural insights considering the development of the route selection space over the first 10 trips of the day are detailed in Section 5.5.1. Findings related to temporal characteristics of bicycle route patterns of tourists are presented in Section 5.5.2. Descriptive statistics of the explanatory variables are discussed in Section 5.5.3. Finally, we elaborate on the correlations among explanatory variables and route selection space determinants.

5.5.1 Route selection space and its dynamics

Routes may deviate to some extent from the bearing line (direct distance between origin and destination) due to the underlying street network, (spatial) knowledge of the environment, and personal preferences such as time constraints. Here, we analyse the route selection space in relation to i.) acquisition of spatial environmental knowledge with each bicycle trip of a tourist undertakes, and ii.) the opportunities that the urban street network provides, which varies with the Euclidean direct distance between origin and destination.

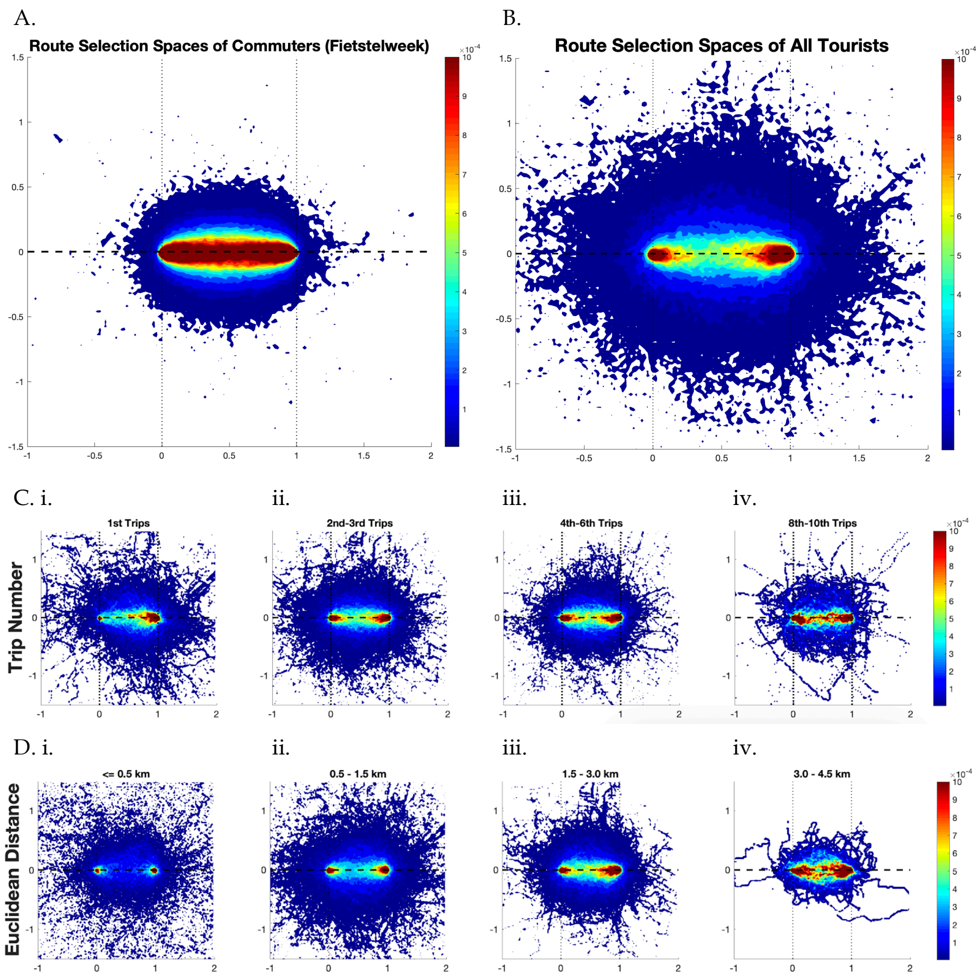


Figure 25 A-D. A. All normalized trajectories plotted with a transparency to visualise the contours of the route selection space. B. Spatial probability plot of the 3 kmx 3 km area around the normalized direct distance between all origins and all destination. The color identifies the probability that a trajectory will pass a cell sized 0,02 x 0,02. C. Subset of the spatial probability plots based on the trip number. D. Subset of the spatial probability plots based on the Euclidean direct distance of the trip.

The spatial probability distribution has previously been used to visualize habitual car trips in four cities by Lima et al. (2016). Therefore, a comparison between habitual bicycle trips in the same urban street network is required to draw conclusions on the route selection space of tourists by bicycle. Figure 25 A shows the spatial probability distribution of bicycle trips in Amsterdam collected during the Fietstelweek (Ton et al. 2017). This dataset consists of habitual bicycle trips of local residents (likely to be commuters) in Amsterdam in 2017. For this figure all GPS trajectories are normalized to geodesic coordinates. We can observe that the dimension of the route selection space is very similar to the habitual trips by car found by Lima et al. (2017). Habitual trips by bicycle in Amsterdam are more likely to stay closer to the bearing line between origin and destination, indicated by the dark red area. The figure also implies no preference in habitual trips for the right or left side of the bearing line.

Figure 25 B shows the high probability region of the route selection space of tourists travelling by bicycle, surrounding the origins and destinations in normalized geodesic space. This figure also illustrates that detours and deviations of tourists are in most cases larger compared to habitual bicycle trips. Moreover, tourists are inclined to only follow the bearing line near the origin and destination. Although the inner shape is similar, the probability within the route selection space of tourists is much lower compared to findings from habitual car trips reported by Lima et al. (2016).

Figure 25B is a compilation of various types of movements, which hides the development of the route selection space over time. Figure 25 C.i-iv illustrates the development of the route selection space with every trip. Here, Figure 25 C.i is a subset of all the first trips of all tourists. These four figures demonstrate a wide dispersion of the first trips of tourists and the higher probability regions increase gradually in strength to become more concentrated around the bearing line with each additional trip that is being performed by the tourists.

In comparison to the habitual bicycle route selection space of commuters, Figure 25 C.iv shows most similarities. This can imply that by 8 trips tourists learn a wayfinding strategy that enables them to find the way more or less similar to the strategy of commuters. Another explanation could be that only tourists that are familiar with the city visit 8 or more activities. A definite answer requires further research and possibly controlled experiments.

Another manner to divide the total set of routes into subsets is by the direct distance between origin and destination. Figure 25 D.i-iv show the differences in the route selection space for varying Euclidean distances between origins and destinations. These figures demonstrate a similar trend as the figures featuring subsets based on trip number. Additionally, Figure 25 D.i-ii show the effect of the underlying urban street pattern on the route selection space. Destinations within a range of 500 metres often require large deviations from the bearing line, possibly due to the width of building facades and limited number of streets. A different explanation pertains to data collection: in some cases the large deviation might be the result of temporal gaps in the GPS data (compromising the quality of activity identification and cleaning the followed path) and tours (when a destination is close to the origin the detour ratio yields extreme values). The low probability ellipse is rather concise compared to destinations in a range of 500 metres to 1.5 km, although there is a significant amount of scatter reaching the border of the 3x3 area. Destinations within a range of 3 to 4,5 km usually extend beyond city boundaries and are characterized by a high probability to remain close to the bearing line near the origin and destination, mid-way tourists stay either close to the left or right side of the bearing line. The combination of these findings also imply that the model needs to control for a possible interaction effect between trip number and Euclidean direct distance.

5.5.2 Trends in the indicators of the route selection space

The route selection space is the result of different travel behaviour preferences. In this chapter four indicators are investigated, namely the detour ratio, the maximum deviation from bearing line, the eccentricity, and the curvature (Table 13).

Table 13. Descriptive statistics of route efficiency indicators.

Metric (N =8.490)	**min**	**max**	**mean (std.)**
Detour ratio	1,00	1.636,20	6,39 (53,30)
Maximum deviation from bearing	0,00	711,99	1,45 (15,97)
Eccentricity	0,00	1,00	0,75 (0,24)
Curvature	0,12	1,00	0,79 (0,15)

5.5.2.1 Descriptive statistics

When the travel distance of a trip equals the direct distance the detour ratio is at the minimum level of 1,00. A ratio of 2,00 implies that the travel distance is double the length of the direct distance. Although the average detour ratio is 6,39, 70% (resp. 90%) of the trips have a detour ratio lower than 2 (resp. 4). Although the average maximum deviation is 1,45, 90% of all trips have a deviation smaller than 1. This implies that 90% of the trips do not deviate by more than the Euclidean direct distance between origin and destination from the bearing line. The mean eccentricity of 0,75 with a standard deviation of 0,24, and the spatial probability density function demonstrates that tourists' routing behaviour is quite eccentric. The curvature represents the detour ratio from the minimum distance to reach the boundary of the route selection space of the respective trip (focal distance). The curvature is 1 when the travel distance equals the focal distance. If the travel distance is twice as long compared to the focal distance the curvature is 0,5. The average curvature of 0,79, with a standard deviation of 0,15 indicating that deviation from the focal distance is rare.

5.5.2.2 Development of the route selection space indicators

Boxplots are used to identify the development trend of the detour ratio, maximum deviation from the bearing line, eccentricity, and curvature (Figure 26 A-D). Values of 1 for detour ratio, eccentricity and curvature correspond to "efficient" movement, and a value of 0 for maximum deviation. Striking in all four boxplots is the stability of the median. The results suggest that detour ratio has the strongest development as 50% of the tourists have on average a decrease of 0,5 after the first four trips and the variation decreases by 1,0. A more moderate decrease continues until the 7th trip. Similar trends can be observed for the maximum deviation and eccentricity. Only the curvature does not show a strong learning curve; the 25% around the median remains constant, and the variation gradually decreases from the fourth trip. The results indicate that while on average there might be little to no improvement ,the percentiles indicate also strong spatial learning effects from the first trip, while it takes a bit more time to find routes with eccentricity and curvature values closer to 1.

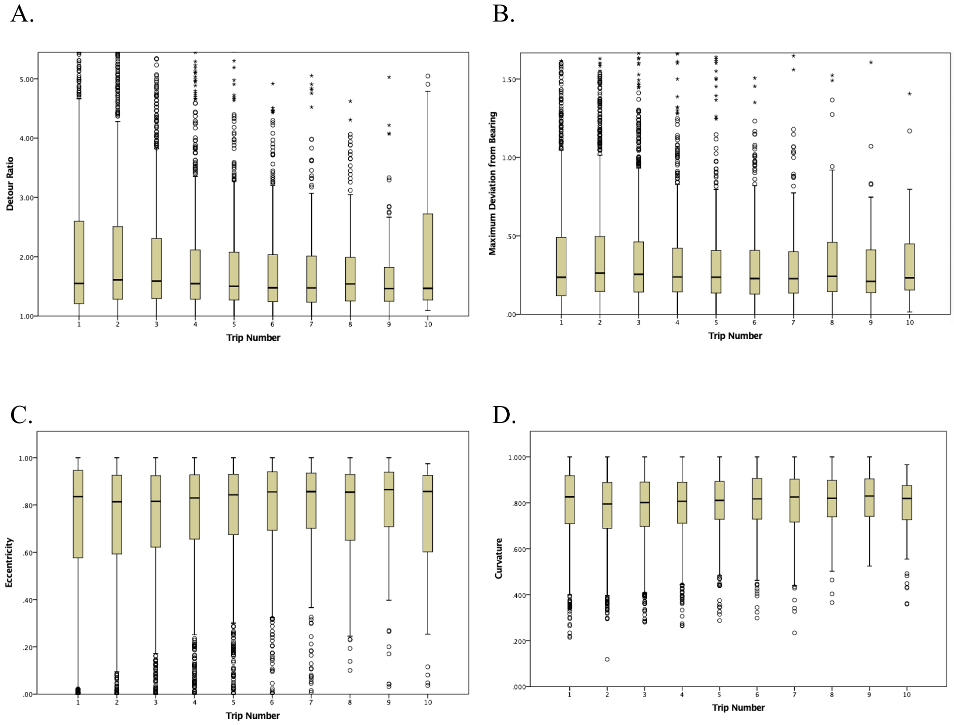


Figure 26 A-D. Development trends of route selection space indicators.

5.5.3 Descriptives of explanatory variables

Most tourists visit 5 to 6 activities by bicycle, with 75% visiting 3 to 8 activities. The direct distance between activities is on average 1,64 km, with 50% of the destinations at a distance between 0,61 and 2,26 km. On average they stay 15 minutes at the destination, with a median travel time ratio of 0,61. This implies that tourists are willing to cycle on average 20 minutes for 10 minutes of activity. The variation in travel time ratio is high, 25% around the median has a travel time ratio between 0,23 and 0,93.

5.5.3.1 Travel experience

The average experienced travel distance is 7-8 km per day. The average bicycle duration is 2 hours and 11 minutes, with a variation of 3 hours and 3 minutes. The average activity duration is slightly longer with 2 hours and 46 minutes.

5.5.3.2 Acquisition of spatial knowledge

Acquired spatial knowledge is measured through 5 determinants. The maximum covered space by a tourist is 13,01 km^2, but the average is almost 10 times smaller. The retrace space is the summation of the surface area that has been visited at least two times during previous activities and trips. The maximum retraced area is 6,88 km^2, but on average tourists' retraced area is 0,24 km^2, smaller than the area induced by a single activity.

The number of anchor areas describes the fragmentation of the retraced area, as we assume that retraced areas have a bigger influence on the (structure of) memorized space. The most fragmented memory of retraced space consists of 11 anchor areas, while 50% of the trips are based on 0 to 2 anchor areas. The last two determinants that describe the acquisition of spatial knowledge are the number of new and recurrent performed activities upon embarking on a trip. When a destination buffer has an overlap with the covered area, it will be counted as a recurrent activity for the next trip. Table 14 shows that the maximum number of new activities (previous destinations without an overlap with the covered area at that time) is 9, while the maximum number of recurrent activities is 23. On average almost 2 new activities and 1,5 recurrent activities have been visited upon embarking on a new trip. Looking at 50% of the trips, there are 1 to 3 new activities and 0 to 2 recurrent activities.

Table 14. Descriptive statistics of the evolvement of dynamics in tourist bicycle travel patterns.

	Determinant	**min**	**max**	**mean (std.)**
Base	Euclidean direct distance	0,00	23,25	1,64 (1,53)
	Trip number	1	29	3,52 (2,5)
Activity/Trip level	Number of activities in a day	2	29	5,69 (2,51)
	Recurrent [0] or new destination [1]	0	1	0*
	Travel time ratio per trip	0,00	1,00	0,7 (0,35)
	Activity sequence	0,03	1,00	0,61 (0,29)
Experience	Experienced travel distance	0,00 km	64,83 km	7,81 (8,43)
	Experienced travel duration	0 min	1.335 min	131,45 (183,46)
	Experienced activity duration	0 min	1.383 min	166,67 (227,84)
Knowledge	Covered space	$0{,}28^{e6}$	$13{,}01^{e6}$	$1{,}77^{e6}$ ($1{,}48^{e6}$)
	Retrace space	0	$6{,}88^{e6}$	$0{,}24^{e6}$ ($0{,}49^{e6}$)
	Number of anchor areas	0	11	0,99 (1,63)
	Number of new activities	0	9	1,98 (1,27)
	Number of recurrent activities	0	23	1,54 (1,74)
Movement	Previous median bicycle speed	0,00 km/hr	25,17 km/h	8,02 (5,11)
	Space variation index	0,00	1,85	0,09 (0,16)

* mode

5.5.3.3 Movement patterns

Finally, the movement pattern determinants consist of median travel speed of the previous trips and the space variation index. The average median bicycle speed is 8 km/h , with 50% of the trips between 3,14 and 11,97 km/h. Median bicycle speeds higher than 13,5 km/h are rare. The space variation index is 0,00 if all areas are visited only once (extreme exploration) while values above 1,00 indicate that the retraced area is larger than the covered area (extreme retracing). According to the LUCY study 60% of the tourist trips are considered “extreme exploration” with a space variation index of 0. Followed by 15% with a space variation index up to 0,12, and values higher than 0,29 being rare, as could be expected from tourists' movements.

5.5.4 Correlations among explanatory variables

Statistical tests are performed to determine whether correlations exist among the explanatory variables, and between the explanatory variables and the dependent variables. Kruskal-Wallis (and Spearman for nominal variables) tests are performed to identify i.) highly correlated explanatory variables, and ii.) significant correlations with the four dependent variables (Table 15).

Table 15. Overview of correlations.

	Determinant	detour ratio	max.deviation	eccentricity	curvature
Base	Trip number	-0,03**	-0,02**	0,03**	ns
	Euclidean direct distance	-0,10**	-0,15**	0,18**	-0,09**
Activity/Trip level	Day movement pattern	-0,07**	-0,05**	0,06**	0,06**
	New destination [ref:1]	0,05**	0,07**	0,08**	0,04**
	Trip Travel Time Ratio	0,17**	0,14**	-0,13**	-0,18**
	Trip Activity Sequence	0,02**	0,02**	ns	-0,05**
Experience	Exp. bicycle distance	-0,03*	-0,02**	0,03*	ns
	Exp. bicycle duration	-0,02*	-0,01**	0,02*	ns
	Exp. activity duration	-0,05*	-0,04**	0,05**	0,02*
Knowledge	Covered area	-0,03**	-0,02**	-0,03**	ns
	Retraced area	-0,04**	-0,02**	-0,03**	-0,02**
	Nr of Anchor areas	-0,04**	-0,03**	-0,03**	-0,02**
	Nr of new destinations	-0,04**	-0,04**	0,06**	-0,03**
	Nr of rec. destinations	ns	-0,02**	0,02**	0,03**
Movement	Median speed	-0,02**	ns	0,02**	ns
	Space variation index	-0,04**	-0,02**	-0,03**	0,02**

** significant at the 0.01 level, * significant at the 0.05 level. ns excluded due to insignificance.

Five explanatory variables show similar effect sizes with all four dependent variables (day movement pattern (e.g. total number of activities in a day), experienced activity duration, retraced area, number of anchor areas, and space variation index). Travel time ratio and euclidean distance to the destination show for all four dependent variables higher correlations.

There are also relative differences between the dependent variables. *Detour ratio* is best explained by day movement pattern, retraced area, number of anchor areas, and space variation index. Only the number of recurrent destinations visited so far does not exhibit a significant relation. *Maximum deviation* scores well on the correlation with euclidean distance, new or recurrent destination, travel time ratio to the destination, but not relatively better compared to the other dependent variables. The median travel speed of all previous trips does not exhibit a significant relation. *Eccentricity* shows the strongest correlations with euclidean distance, new or recurrent destination, and number of new destinations visited so far. But the travel time ratio to the destination has the lowest effect, compared to the other dependent variables. *Curvature* has less, and often the weakest, significant relations with the explanatory variables, two exemptions are travel time ratio of the trip to be estimated and the activity sequence (position of the destination in the complete activity chain).

The results of the statistical tests are used to determine which variables should be included in the GEE model. Here, explanatory variables with the standardized correlation coefficient is lower (higher) than -0,8 (0,8) should not be included simultaneously in the same model. In the case of strong correlations between variables, the variable with the largest effect on the dependent variable is included if both explanatory variables are significantly related with a dependent variable ($p < 0,05$). The following three groups of explanatory variables show high correlations: 1.) experienced bicycle distance with covered area, 2.) anchor points, retraced area, and space variation index, and 3.) trip number with experienced bicycle distance, covered area, retraced area, anchor points, and recurrent destinations. All significant effect sizes are weak.

5.6 Modelling the route selection space

The Generalized Estimating Equations (GEEs) have been estimated based on the robust covariance matrix, the autoregressive 1 (AR1) working correlation matrix, and assuming a normal distribution with an identity link function. Therefore, only the relative trends can be interpreted from the model results. Furthermore, we consider all bike records which consist of no more than 29 trips, which accounts for more than 98% of the trips.

The results of the GEEs are described by route selection space characteristics. The detour ratio is discussed in Section 5.6.1, the maximum deviation in Section 5.6.2, the results related to eccentricity are presented in Section 5.6.3, and the elaboration on the curvature can be found in Section 5.6.4 (Table 16). This chapter finalizes with a synthesis on the model results, spatial learning, and urban tourism management (5.6.5).

Table 16. Model results of robust GEEs with AR1, normal distribution & identity link function.

N = 3.350	**Detour ratio**		**Max. deviation**		**Eccentricity**		**Curvature**	
Parameter	B	Std. Error	B	Std. Error	B	Std. Error	B	Std. Error
(Intercept)	46,31	8,36**	6,21	1,66**	0,78	0,01**	0,85	0,01**
Trip number	-3,46	1,12**	0,01	0,35	0,00	0,00	-0,03	0,00**
Euclidean direct distance	-3,94	0,73**	-1,02	0,22**	0,05	0,01**		ns
Day movement pattern	-1,41	0,54**	-0,40	0,17*		ns	0,01	0,00**
New destination [ref:1]	26,24	4,49**	5,98	1,36**	-0,12	0,01**		ns
Trip Travel Time Ratio	20,28	4,46**	5,33	1,51**	-0,30	0,01**	-0,16	0,01**
Trip Activity Sequence	-18,69	7,72*	-8,13	3,45*			0,03	0,02
Exp. bicycle distance								
Exp. bicycle duration	-0,02	0,00**			0,00	0,00**		
Exp. activity duration		ns	0,00	0,00*	0,00	0,00**	0,00	0,00**
Covered area								
Retraced area								
Nr of Anchor areas								
Nr of new destinations	-4,80	1,44**	-1,70	0,48**		ns	-0,01	0,00**
Nr of rec. destinations								ns
Median speed	-0,83	0,28**			0,01	0,00**		
Space variation index				ns				ns
Interaction effects with Trip number								
* Euc. direct distance	0,41	0,13**	0,11	0,04**	0,00	0,00		
* New destination [ref:1]	-4,85	0,88**	-1,12	0,27**	0,02	0,00**		ns
* Trip Travel Time Ratio	-2,30	0,74**	-0,64	0,27*	0,03	0,00**	0,01	0,00**
* Trip Activity Sequence		ns	0,76	0,46			0,02	0,00**
* Experienced bicycle duration	0,00	0,00**			0,00	0,00*	0,00	0,00**
* Experienced activity duration				ns	0,00	0,00		
* Nr of new destinations	0,67	0,20**	0,27	0,07**		ns		ns
* Median speed	0,21	0,07**			0,00	0,00*		

** Significant at the 0,01 level, * Significant at the 0,05 level. Grey: not included due to correlation >0,8. ns: Excluded due to insignificance.

5.6.1 Detour ratio

The determinants with the most positive and negative factor loadings for the detour ratio are part of the trip and activity characteristics, namely whether it is a new or recurrent destination and the total number of activities that will be visited during the day. The model illustrates that if the destination of the trip is in an unexplored environment the detour ratio increases. However, the interaction effect between trip number and a new destination also reveals a learning effect, whose effect decreases with each additional trip.

Ultimately, if a new destination is visited after the sixth trip the detour ratio is likely to be smaller compared to visiting a recurrent destination after the sixth trip. The activity sequence of the trip provides concerns how many trips have been visited and how many will still be visited. The model reveals one other main effect, namely trips further down the sequence closer to the final destination correspond with smaller detour ratios. There is no interaction effect with how many trips have been made.

In addition, Euclidean direct distance has a (relatively small) main effect on the detour ratio of tourists by bicycle, while controlling for the number of trips. The further apart two activities are, the larger the detour ratio in general is. However, similar to visiting a new or recurrent destination, there is a learning effect with every trip, which makes the detour ratio independent from the Euclidean direct distance, after 9 to 10 trips. As the model is limited to the 10th trip, a longitudinal dataset can possibly reveal at which number of trips the Euclidean distance becomes insignificant.

5.6.2 Maximum deviation from bearing

The determinants with the positive and negative factor loadings for the maximum deviation from bearing line are also part of the trip and activity characteristics, namely the total number of activities that will be visited that day, new or recurrent destination, and travel time ratio of the trip. The model illustrates that the further down the activity sequence and closer to the final destination of the day, the smaller the maximum deviation becomes with respect to the bearing line. Near the end of the activity sequence, the effect on the deviation from the bearing line is, on average, 0. Here, the model also establishes a negative learning effect, being the more trips already have been made, the smaller the negative effect of the length of the activity sequence on the maximum deviation. Besides that, new or recurrent destination and travel time ratio of the trip are both positively related to the maximum deviation. Also in this case, a positive learning effect is found; after 5 trips a trip to a new destination , visiting a new destination does not correspond to a larger deviation ($5{,}98 + (5 \times -1{,}12) \sim 0$). Considering the average ratio between travel times and activity duration (0,61), only after 14 trips the effect of the travel time ratio on the deviation is neutralized.

Last of all, similar to the detour ratio, the Euclidean direct distance has a negative relation with the maximum deviation from the bearing line, which also neutralizes due to a learning effect after 9 to 10 trips.

5.6.3 Eccentricity

The third model reveals that the determinants with the most positive and negative factor loadings for the eccentricity are the travel time ratio and the Euclidean direct distance. There is a negative main effect of travel time ratio, indicating that trips become more eccentric when travel times become longer compared to the activity duration, i.e. when the trip itself becomes an activity too. However, there exists a positive learning effect, which identifies that when more trips have been made, higher travel time ratios will increase the centricity. The model also reveals that larger Euclidean direct distance increases the centricity. However, there is no learning effect found between Euclidean direct distance and eccentricity, although the interaction effect with the number of trips is significant, the effect size is close to zero.

5.6.4 *Curvature*

The last model reveals that the determinants with the most positive and negative factor loadings for the curvature are travel time ratio and the activity sequence. Higher travel time ratios decrease the curvature index leading to less efficient paths. The model also reveals trips temporarily closer to the final destination of the day on average have an increased curvature index. With increased trip number the activity sequence will increase the positive effect to improve path efficiency. Travel time ratio shows a minor learning effect that will not neutralize the main effect within 10 trips ($-0{,}16 + (10 \times 0{,}01) > 0$). The Euclidean direct distance did not have a significant effect on the curvature and has been excluded.

5.6.5 *Synthesis*

A comparison of the model results brings two findings. When controlling for trip number the route selection space characteristics depend mostly on trip dynamics (new or recurrent destination, travel time ratio, and total number of activities that will be visited in a day). Secondly, the detour ratio and the maximum deviation reveal relatively stronger effects with determinants. The detour ratio is relatively well described by the trip number, experienced bicycle duration and median travel speed, while the maximum deviation is relatively well described by the Euclidean direct distance, number of activities in a day, new or recurrent destination, number of new activities, and travel time ratio and activity sequence of the trip.

Regarding spatial knowledge acquisition the space variation index does not reveal a significant relation with the route selection space in any of the models. Another observation relates to the high correlation effects with trip number. The covered area, retrace area, number of anchor areas, and experienced bicycle distance have not been included in any model due to high correlation effects with other explanatory variables that yield a stronger effect on the dependent variable. A possible explanation for both findings could be that the tourists in this chapter have a strong focus on exploring Amsterdam by bicycle, and did not retrace their paths systematically. A longitudinal data set on routing behaviour of tourists would be necessary to identify if and when spatial knowledge acquisition is matured. As spatial knowledge acquisition is matured for habitual trips, this also suggests that acquired spatial knowledge is important for the route selection space of commuting behaviour.

5.7 Conclusions and implications

This chapter focussed on the development of the route selection space when new knowledge is acquired with every bicycle trip of tourists. This was done to answer the main research question of this chapter, which was *"To what extent does experience with the built environment influence the development of spatial behavior of tourists by bicycle over time?"*

Generalized Estimating Equations have been used to assess the learning effect of tourists as a function of the estimated trip purpose, familiarity, and movement patterns. Four route selection space characteristics have been investigated: detour ratio, maximum deviation from the bearing line, eccentricity, and curvature.

This chapter revealed that the spatial density probabilities of the route selection space deflates with every trip and increased Euclidean direct distance between origin and destination. The route selection space characteristics depend mainly on trip purpose (new or recurrent destination, travel time ratio, and the number of activities that will be visited in a day). Contrary to previous findings in literature based on habitual car trips, longer Euclidean direct distances between origin and destination decrease the relative detour and maximum deviation, and increase the centricity. The acquired knowledge of tourists can be captured by the current number of trips and number of new activities that have been visited. Here, it should be noted that the trip number is highly correlated with experienced bicycle distance, covered area, retraced area, anchor points, and recurrent activities. The GEE models also indicate that the effect of learning generally stabilizes after 8 or more trips. Combined, these insights suggest that experienced travel distance, covered area, retrace area, number of anchor points, and number of recurrent activities in a daily travel pattern may influence the route selection space of habitual trips.

These findings can have major implications for route choice modelling in the field of tourism, and beyond. In combination with the street network and origin-destination matrices, the contour probabilities of the route selection space can be used to model the probability of route trajectories. Another application could be to include the route selection space as a selection rule of the link elimination procedure (e.g. first eliminate most likely links).

The results also provide empirical underpinned behavioural insights for the management of urban tourism in Amsterdam in 2030. The forecasts expect 3 to 9 million more tourists to visit Amsterdam in 2030 (UNWTO 2018b; NBTC 2019), while the number of citizens is expected to have a relatively small growth of 0,1 million (PBL/CBS 2016). Good management can disperse tourists by bicycle to outer areas and will not only prevent "overtourism", but it will also economically benefit deprived areas. The results of this chapter indicate that with longer Euclidean direct distance to reach activities in the outer areas, the detour ratio and maximum deviation decrease and trips become more centric. Also if intermediate activities are promoted to cut a long trip in two segments, increased trip number, number of activities that will be performed in a day and the activity sequence also correspond to a decrease in

detour ratio and maximum deviation and increase of centricity. Therefore, it can be concluded that tourists need to be able to follow a bicycle route along a more directed street network to reach outer areas. This does not necessarily entail more streets, as tailored route advice for tourists also provides an opportunity, for example if it takes unique sites into account and allows for a larger detour from the shortest path.

The GEE models' ability to describe the route selection space of tourists travelling by bicycle still requires more testing. The distribution of route selection space metrics suggest GEEs based on a tweedie distribution are more suitable. Estimation of models without trip purpose characteristics can provide insights into the learning effect of spatial knowledge. Additionally, street network characteristics based on space syntax and topological distance can be investigated to test different concepts of spatial knowledge. Finally, the influence of memory decay will be essential to include in future studies to advance the model to describe the development of spatial knowledge. However, we also stress that the presented insights of the development of the route selection space are already useful to design, test, and evaluate their policies, design, and travel information to realize and maintain bikeable cities for citizens, as well as, tourists.

Acknowledgements

The research leading to these results has received funding from the European Research Council under the European Union's Horizon 2020 Framework Programme for Research and Innovation. It is established by the Scientific Council of the ERC Grant Agreement n. 669792 (Allegro).

THE VISITOR GUIDE
BLUE PEARLS of
Delft
LEGEND/LEGENDA
FOOD & DRINK
ETEN & DRINKEN
FASHION
MODE
LIFESTYLE
LIFESTYLE
ATTRACTION
BEZIENSWAARDIGHEID
ATM 24/7 GELDAUTOMAAT
IN STORE ATM
PARKING
PARKEREN
SECURED BICYCLE PARKING
BEVEILIGDE FIETSENSTALLING
HISTORICAL SHOPPING AREA
HISTORISCH WINKELGEBIED
MODERN SHOPPING AREA
MODERN WINKELGEBIED
CAR FREE ZONE
AUTO LUWE ZONE
UNIVERSITY OF DELFT CAMPUS
TU DELFT CAMPUS
HOTEL
BED & BREAKFAST
TRAM STOP
Plattegrond © Pixlbar 2014
MARKT
STATION
TRAIN
BUS
BOTANIC GARDENS
10 min. walk 10 min. lopen

THE VISITOR GUIDE
BLUE PEARLS of
Delft
LEGEND/LEGENDA
FOOD & DRINK
ETEN & DRINKEN
FASHION
MODE
LIFESTYLE
LIFESTYLE
ATTRACTION
BEZIENSWAARDIGHEID
ATM 24/7 GELDAUTOMAAT
IN STORE ATM
PARKING
PARKEREN
SECURED BICYCLE PARKING
BEVEILIGDE FIETSENSTALLING
HISTORICAL SHOPPING AREA
HISTORISCH WINKELGEBIED
MODERN SHOPPING AREA
MODERN WINKELGEBIED
CAR FREE ZONE
AUTO LUWE ZONE
UNIVERSITY OF DELFT CAMPUS
TU DELFT CAMPUS
HOTEL
BED & BREAKFAST
TRAM STOP
Plattegrond © Pixlbar 2014
MARKT
OUDE KERK
BEESTENMARKT
BOTERBRUG
STATION TRAIN
BUS
PHOENIX
PRINSENHOF
PAARDENMARKT
GASTHUISPLEIN
ZUIDPOORT
BOTANIC GARDENS
NIEUWE PLANTAGE
KONINGSPLEIN
KANTOORGRACHT
10 min. walk / 10 min. lopen

6 Conclusion

6.1 Main Contributions and Findings

The overarching objective of the studies in this dissertation is to unravel how travel behaviour, urban environments, and information services relate to spatial knowledge development. The contributions and findings narrowed the gap between research on travel behaviour research and research on urban spatial knowledge (6.1). Based on limitations pertaining to this dissertation and insights we also provide fruitful direction for future research (6.2). In order to develop theory, conceptual and mathematical models on the development of active modes' spatial knowledge in activity-travel modeling across urban environments, the overall research objective was:

Unravel the role of spatial (network) knowledge and how people find their way in urban environments.

The conclusions, in terms of contributions and implications, of are fourfold:

1. We tested a wayfinding theory based on existing literature to identify the components of urban wayfinding behaviour, and investigated the relation with daily travel behaviour based on travel diary data (further detailed in 6.1.1).
2. We developped a methodology based on spatial analytics to use open spatial data to characterize salient and legible areas in an urban environment (further detailed in 6.1.2).
3. We gained insights into activity patterns of tourists based on a large empirical field study to collect GPS trajectories of bicycles. The insights are used to develop new theories to better understand and influence travel behavior of tourists by bicycle in crowded cities (further detailed in 6.1.3).
4. By collecting GPS trajectory data of bicycles of tourists and combining these with spatial maps from open data, we have computed determinants that describe the (development of) spatial knowledge level of tourists. A model is used to estimate to what extent the detour ratio and deviation area of a bicycle

trip can be predicted based on the theoretic acquired spatial knowledge (further detailed in 6.1.4).

6.1.1 *Urban Wayfinding*

Test a theory based on existing literature to identify the components of urban wayfinding behaviour, and investigate the relation with daily travel behaviour basedon travel diary data.
***RQ:* What are the components of urban wayfinding styles, and how do they relate to daily travel behaviour? (Chapter 1)**

In literature orientation ability is commonly referred to as the latent variable that captures three basic types of spatial orientation and spatial knowledge (Stea & Blaut 1973; Siegel & White 1975; Golledge & Gärling 2001). However, the classification based on literature is not mutually exclusive, i.e. one can simultaneously rely on egocentric and fixed-point orientation. To identify the components of urban wayfindings styles an exploratory factor analysis is used to transform the latent variables into unique wayfinding styles based on the Santa Barbara Sense-of-Direction (SBSOD) self-report questionnaire originally developed by Hegarty et al. (2002) (*Chapter 1*). Through 23 questions insights are gained on the attitudes towards spatial knowledge acquisition (exploration), orientation within an environment, mental representation of the environment, anxiety, and usage of route information. All respondents are asked to indicate how much a statement reflects their behaviour, ability, or attitude. The most applicable (consistent) number of components is two, including 19 out of 23 questions. The two components are coined *Orientation Ability* (attitude and basic skills to be able to orient and navigate effectively in an urban environment) and *Knowledge Gathering & Processing Ability* (attitude and preferences to extend knowledge about the environment, e.g. explore cities and take new routes). Each wayfinding style relies largely on unique variables, while also three common variables exist: ability to give route directions, perception of distances, and attitude to read maps. These results advocate that Orientation Ability (OA) and Knowledge Gathering & Processing Ability (KA) are partially dissociated. Categorization of the latent variables to three levels with cut-off values at −0,5 and 0,5 transforms the latent variables to three categorical wayfinding styles per component: lower than average (-1), average (0), and higher than average (1).

Two holistic Generalized Linear Models describe urban wayfinding styles based on both categorical factor components “Orientation Ability” and “Knowledge Gathering & Processing Ability”. The model results are acceptable as they are a 42% improvement compared to a random accuracy of 0,33. However, the quality of the model results is not equally distributed; “low” and “high” wayfinding abilities are better modelled compared to the respective “average” wayfinding ability. The following determinants are significant: gender, age, education level, perceived bicycle accessibility of the neighbourhood, attraction to familiar and unfamiliar streets, and greenopy of the streets, navigational preferences to minimize turns, follow the bearing line, and take short-cuts, ratio of active navigation, average daily distance travelled by car and bicycle, and average daily number of trips made by car. Gender and age have similar effect signs on both OA and KA, while the navigational preference to follow the bearing line and average daily distance travelled by car have disassociated effects. The remaining determinants are only significant in either OA or KA, providing evidence that predominantly different processes describe each urban wayfinding style.

From the literature background it was hypothesized that the total average travel distance (by car and foot) have a negative relation with the wayfinding score. The results in this chapter also show a negative relation for distance travelled by car, and for the first time, also distance bicycled by men with OA. However, this chapter also shows that the total average distance travelled by car and the interaction effect between average number of car trips and active navigation ratio have positive relations with KA. Although the majority of the research found in literature investigates pedestrian wayfinding, the distance travelled by foot and public transport are not significant in this chapter.

Based on the models it seems that a combination of high OA and KA will correspond to higher variability in the streets of chosen routes. With higher (perceived) connectivity of the bicycle infrastructure more OA is required than average. This implies that people with lower levels of OA will compensate for the complexity of the urban wayfinding task by preferring a longer route along familiar streets. Thus, even if high connectivity exists, but all people have low orientation abilities, still not much route variation will occur and it will become more difficult to mitigate congestion and distribute large cyclists flows more evenly. Insights related to navigational preferences and urban environment on KA can be interpreted as for people that do not wish to make short-cuts, for example due to absent time pressure, it is easier to memorize a detour through a green passage. Last, although urban density has been identified as important characteristic for salience and legibility of an environment, its role as a determinant remains unknown, as neither model indicated significance.

Both wayfinding styles can be used complementary as different processes influence them. However, two determinants (navigational preference to follow the bearing line and average daily distance travelled by car) have an ambiguous effect on both wayfinding styles. This could indicate a trade-off, because gathering and processing more spatial knowledge will ultimately require more orientation ability in order to process the knowledge into useable wayfinding styles. The navigational preference to follow the bearing line is not beneficial when there is a low amount of spatial knowledge, as this does not encourage the acquisition of more spatial knowledge. If a satisfactory amount of spatial knowledge has been acquired using the bearing line as a navigational preference is useful to reduce the workload.

6.1.2 Salient Urban Areas

Develop a methodology based on spatial analytics to use open spatial data to characterize salient and legible areas in an urban environment.

How can open spatial data be used to identify salient and legible urban areas (landmarks)? (chapter 2)

City users, to some extent, rely on memorized urban route knowledge to decide how to move from one place to the next. To this end, spatial urban route knowledge can be viewed as remembered sequences of landmarks, that, combined with directional actions support users to navigate across town. Following Lynch (1960) and Appleyard (1970), landmarks are defined as salient geographic objects, points, or polygons of buildings that structure the internal representation of a city (Richter and Winter 2014). Determining the location of distinctive landmarks is thus important in research on route choice, urban cognition, and travel information.

Over the last two decades, different approaches to identify and integrate landmarks have been developed, as can be noticed, e.g. in route descriptions. Currently, most approaches to identify landmarks require vast data collection efforts (Richter and Winter 2014). Consequently, knowledge on the effects of urban landmark distribution on wayfinding behavior remains limited. To overcome these demands, this chapter proposes a spatial analytic method able to process and analyse open-source datasets to identify urban wayfinding landmarks as salient urban areas.

The proposed method combines insights from cognitive sciences and spatial analytics from urban morphologies to identify aggregated local and global urban landmarks based on salient characteristics (*Chapter 2*). Also, the concept of aggregate urban landmarks, coined as salient urban areas, is introduced. Salient urban areas possess noticeable characteristics that make them distinct from their surroundings. From a theoretical perspective, a landmark is salient (distinct) in relation to its immediate surrounding or context at large. Salient urban areas are considered unique, either because of dissimilarities to their (local) area, and/or else, because of characteristics considered similar in comparison to other (global) areas. Presumably, the more distinctive a landmark or area, the easier it will be to memorize and incorporate this saliency into the spatial route knowledge to be drawn upon in future. Therefore, salient urban areas are hypothesized to be important to structure spatial knowledge in long-term memory (Couclelis et al. 1987; Sadalla et al. 1980; Montello 1997).

The spatial analytic method consists of five steps based on data management, grouping analysis, and cluster and outlier analysis. Determinants to identify salient urban areas are building volume, surface, height, building year, and the number of buildings in a 100 square meters grid-cell. Results have been applied to identify the differences in distribution of cluster and dispersion between local and global salient urban areas using the Gini coefficient, based on an open-source GIS dataset on the built environment of Amsterdam.

Implications of identifying salient urban areas can provide new insights on how to analyze how wayfinding landmarks structure environmental knowledge and investigate influences on wayfinding strategies. This environmental knowledge (configuration of landmarks) is assumed to become available when also knowledge has been memorized about the general interrelationships between landmarks (Hirtle and Hudson 1991). If people use these wayfinding landmarks as part of the wayfinding strategy, this is expected to be observable in their route choice behavior. For example it could be more likely to take a detour if more wayfinding landmarks will be passed.

6.1.3 Activity Patterns

Gain insights into activity patterns of tourists based on a large empirical field study to collect GPS trajectories of bicycles. The insights are used to develop new theories to better understand and influence travel behavior of tourists by bicycle in crowded cities.

What is the relation between spatial and temporal activity patterns of tourists by bicycle in the metropolitan area of Amsterdam? (chapter 3)

Metrics and methods have been developepd that can be used to characterise the spatiotemporal travel and activity patterns of tourists with access to a bicycle in metropolitan areas (*Chapter 3*). The analyses are based on GPS data of tourists' cycling behaviour collected in the Amsterdam Metropolitan Region in July and August 2017.

Data processing approach that was used in this chapter classified stationary and moving GPS points that have been collected during July 1st and September 1st, yielding 10.342 activity locations and 8.525 trips, made by 1.817 unique tourist days. This information has been used in four analyses to unravel spatial and temporal travel and activity patterns of tourists: k-means clustering identified 105 activity zones, network analysis identified spatial relations between activity zones based on four communities and betweenness centrality of the Top 15 most visited activity zones, activity space analysis investigated the spatial dispersal of tourists, and temporal profiles identified which moments of the day most tourists are visiting activities or bicycling in and around Amsterdam.

Compared to pedestrian studies on urban tourism (Shoval and Isaacson 2007; Van Der Spek 2010), our findings indicate that it is possible to achieve wider spatial dispersal of tourists if they travel by bicycle, provided that activities in outer areas are promoted at hotels, chapter and online tourists maps are available with the bicycle infrastructure and a clear overview which activities are feasible to visit by bicycle during one, two, or three days. The results also suggest wayfinding systems at A-locations indicating bicycle times for tourists, if they are lower than the expectation of tourists. If the outer areas belong to different (same) activity communities, the locations with high betweenness centrality (clustering coefficient) have the most chance to disperse tourists to other communities. To alleviate local crowdedness, activity locations within a radius of 2-3 kilometres from the most crowded sites, are potential spill-over zones. Secondly, the location of hotels appears to influence the activity pattern and travel behaviour of tourists travelling by bicycle. Moving hotels closer to the A10 could increase the visits at less common activities within the 2-3 kilometre radius.

Other insights pertain to the activity network of tourists with access to a bicycle, which appears to be consistent with expectations from other transport networks; there are differences between communities, and correlations with betweenness centrality are positive for closeness centrality and weighted degree, while negative for the clustering coefficient. Secondly, the combination of activity spaces, travel ratios, and travel pattern structure extend existing theories of tourists' travel behaviour (e.g. McKercher & Lau 2008). Mobility patterns of tourists using bicycles in metropolitan regions vary from activity oriented to trip orientated, and from directed to compact. Further research is required to explore the difference between citizens and tourists, and tourists travelling by foot and public transport.

6.1.4 Dynamics of the Route Choice Ellipse

GPS trajectory data of bicycles of tourists is combined with spatial maps from open data to compute determinants that describe the (development of) spatial knowledge level of tourists. A model is used to estimate to what extent the detour ratio and deviation area of a bicycle trip can be predicted based on the theoretic acquired spatial knowledge.

What determines the spatial boundary of the route selection space of tourists travelling by bicycle, and how does spatial (network) knowledge acquisition influence the movement pattern to the next activity? (chapter 4)

The spatial probability distribution is used to visualize the route selection space (*Chapter 4*). Although some routes have very large deviations, most deviations are relatively small. Detours and deviations are in most cases well-bounded. Moreover, tourists are inclined to follow the bearing line near the origin and destination. Although the shape is similar, the probability within the route selection space of tourists is much lower compared findings of habitual car trips reported by Lima et al. (2016).

The spatial density probabilities of the route selection space contracts with every trip and increased Euclidean direct distance between origin and destination. The route selection space characteristics depend mainly on trip purpose (new or recurrent destination, travel time ratio, and the number of activities that will be visited in a day). Contrary to previous findings in literature based on habitual car trips, longer Euclidean direct distances between origin and destination decrease the relative detour and maximum deviation, and increase the centricity. The acquired knowledge of tourists can be captured by the current number of trips and number of new activities that have been visited. Here, it should be noted that the trip number is highly correlated with experienced bicycle distance, covered area, retraced area, anchor points, and recurrent activities. The Generalized Estimating Equation models also indicate that most learning effects stabilizes after 8 or more trips around the median. Combined, these insights suggest that experienced travel distance, covered area, retrace area, number of anchor points, and number of recurrent activities in a daily travel pattern may influence the route selection space of habitual trips.

6.2 Limitations & Future Work

The limitations and future work pertain to wayfinding, travel information, perception, network, route choices, and the relation between science and practice.

6.2.1 Evolvement of urban wayfinding styles

To what extent and how do urban wayfinding styles evolve in time?

The study on the determinants of urban wayfinding styles distinghuises different orientation and knowledge processing abilities related to travel behaviour based on a special edition attached to the MPN (Dutch Mobility Panel) survey. If the special issue is conducted more frequently in combination with the MPN, changes in wayfinding styles can be investigated, while controlling for life changing events such as home and/or work relocation to a new city. Currently, there are indications that wayfinding abilities develop during childhood, however the development in later stages in life has not been systematically investigated. These insights are essential as societies and urban landscapes are changing due to, amongst others, urbanization and technological advancements.

6.2.2 Influence of emotion in wayfinding

Together with anger, disgust, fear, sadness and surprise, happiness is considered a basic emotion (Ekman & Friesen 2013). Advances in measurement methods enabled economics of happiness to identify relations with GDP, leisure and freedom (Graham 2008). For example mobile apps and questionnaires investigated the role of emotional states to find the way through cities to improve cartography and urban spaces (Gartner 2012; Lee et al. 2017). In a VR experiment other researchers found that motivation can improve wayfinding performance in all but the most complex conditions (Srinivas 2010). Due to safeguarding privacy and the large-scale set-up of the LUCY study, it would be too invasive to combine activity patterns and routing trajectories with other experiments and surveys (Klasnja et al 2009).

Even more recent advances pave the way for large-scale emotion detection using smart cameras. The Artifical Intelligence field already contains numerous examples of facial recognition where not only a smile, but also gender, age, distress, and even race can be identified (Cohen et al. 2003). Future work will be focused on simulteanous detection of the number and happiness levels of cyclists from video feeds. Insights are expected to increase the understanding of the intrinsic motivation and possibilities for long term behavioural change. Understanding the cyclist mood enables policy makers to identify which intersections or road segments have a positive impact on cyclists.

6.2.3 Travellers' information

Controlled experiments are required to investigate how travel information should be adapted to stimulate spatial knowledge development and increase route pattern efficiency. Especially, the role of (familiarity with) the urban environment, orientation ability and knowledge gathering and processing ability need to be taken into account. Moreover, the role of salient urban areas on routing patterns can enhance the quality of travel information when tailored to archetypical travellers. The next step is to identify the effectiveness of travel information to disperse travellers to activity areas outside the crowded city centres.

6.2.4 Perception and influences of the built and natural environment

Urban wayfinding behaviour also depends on the cognitive distance between urban spaces caused by a the mental manipulation process (Chastril and Warren 2013). To facilitate learning while moving along a route information has to be clustered. For forthcoming trips, decomposed information is retrieved from memory consisting of past experiences of the sequences of perceived events (Atkinson and Shiffrin 1968). Upon decision-making along the new route, further distortion may be triggered due to mistakes in reassembling, e.g. distances, directions, sequences of events, or (locational) information of buildings (Golledge, Dougherty and 1995; Ruddle, Volkova and Bülthoff 2011). One of the issues still under discussion, is the way in which our mental reconstruction of the urban environment differs from the actual urban network structure, and how does this translate into difference in movement patterns?

A preliminary study investigated several factors based on 12 distance estimations per respondent at a ratio scale and applied on a case study in the historical city center of Delft, The Netherlands May 2016. The results indicate that at short distances 75% of the responses were overestimated the direct distances and participants who frequently cycle have overall a longer perception of direct distances.

Related questions that are still unanswered are to what extend landmarks along a route also serve as waypoints, or do they only activate certain decision-making processes? Secondly, to what extent can landmarks mitigate errors and distortions of spatial knowledge by resetting the waypoints coordinates? Thirdly, to what extend are different types of landmarks isolated in different fashion (Lynch 1958; Collett and Graham 2010)? The outcomes of the four studies included in this dissertation ask for a comparison of route trajectories (route deviation and utilization of urban salient areas) between tourists, commuters, and travel information provision (e.g. Google Directions).

6.2.5 *Influences of the urban network*

Route and/or network structure analysis can be applied to represent, analyze and characterize to generate new expressions of typologies as well as being used to differentiate types of network, based on their structure of routes. Route structure analysis is based on the contention that the structure of a network is a product of the way that the routes connect up with each other. Route structure analysis is built on three basic route properties: continuity, connectivity and depth. Essentially it adds a third dimension (continuity) to the space syntax method. Combined the three properties represent the routegram. The routegram may be used to compare the structural role of different routes in a network, or different routes across networks, or to compare the complete set of routes in a network with those in another network.

As routes are only a part of the travel pattern, a weighted average of all the points on the routegram belonging to a single person may be used to represent a combination of the relative continuity, connectectivity, and depth for a complete tour in a *tourgram* (also known as *netgram*). Just as it was possible to construct the netgram as a triangular plot from three relative properties summing to one, it is also possible to construct a triangular plot from three of the differentiation properties, namely, complexity, regularity and recursivity to identify different kinds of heterogeneity. The resulting plot would be the *hetgram*, since it addresses the issue of heterogeneity, assisting the recognition of networks according to the differentiation of route types.

6.2.6 *Modelling urban route choices while choice-set is under development*

Discrete choice paradigm constitutes a sound methodology for route choice modelling in relatively simple networks (eg. daily intercity commute on highways). There are several limitations when the network becomes larger and more complex (eg. urban bicycle trips) and the routing behaviour is not aimed at traversing the shortest path (eg. tourism or newcomers). Assigning the probability of each possible alternative is computationaly very complex, and deterministic, stochastic and constrained enumeration methods are all based on shortest path searches (Ton 2019). The route selection space (choice ellipse) provides a behavioural constrain that can be combined with a probabilistic method as proposed by Manski (1977). The spatial probability distribution can be used to assign probabilities to all route trajectories contained within the choice ellipse, and thereby decreasing the computational complexity.

Further research is required to include learning behaviour into the route selection space. This research indicated that learning behaviour yields a dynamic composition of the route selection space indicators, as they evolve with every new trip. For habitual trips transport researchers advocate that people choose a route from a set of considered trips ⊂ known trips ⊂ feasible trips ⊂ logical trips ⊂ all existing trips (Hoogendoorn-Lanser and Van Nes 2004). However, without spatial knowledge the choice sets until known trips are “empty” and instead grow with every new trip. Therefore, it would be fruitful to investigate the possibilities to use decision field theory (DFT) to identify the development of preferences related to the choice set composition. The starting point would be the adapted DFT methodology by Hancock, Hess and Choudhury (2018) that has been applied in various travel choice situations (Hancock, Hess and Choudhury 2018).

6.2.7 Connecting science and practice

The last decade has seen a steady rise in bicycle research (Dill 2017). This rapid development makes the old argument that bicycle research is receiving little attention outdated. Yet, a clear and concise overview to identify (practically) relevant current or new bicycle research does not exist. Consequently, there is a need to synthesize and generalize the most relevant results to a unified and integrated bicycle framework, that helps identify new challenges for bicycle research in the next decade (Zomer et al. 2019).

The impact in the future can identified based on the central themes of cycling research. Questions that help us identify these themes are: Which research efforts have been relevant for the current state-of-the-art, and an inspiration for future? Where should we direct our research efforts to ensure that cities will be more livable, inclusive, and healthy places to live, work, and play in? Do we need more collaboration within academia, or is a stronger connection between academia and practice needed to identify and analyse future bicycle problems (Zomer et al. 2019)?

Preliminary results from an expert-based survey indicated that the importance of generalization of results, which demands a unified framework, standardization in data collection to enable data exchange, consistency in terminology and opportunities to meet and exchange ideas. This requires distinguishment of findings in countries (or cities) with low, high, emerging, and falling bicycle shares. Moreover, the preliminary results suggest problems related to safety, infrastructure and transferability of results can benefit from collaboration within and between fields such as policy, behaviour, attitudes and data. Especially behavioural change in combination with competition among travel modes and implementation barriers is expected to be next challenge to solve (Zomer et al. 2019).

References

ABF Research. (2003). Achtergrond informatie Woonmilieu-typologie.

Ahas, R., Aasa, A., Mark, Ü., Pae, T., & Kull, A. (2007). Seasonal tourism spaces in Estonia: Case study with mobile positioning data. Tourism Management, 28(3).

Anselin, L. "Local Indicators of Spatial Association-LISA," Geographical Analysis 27(2): 93–115, 1995.

Appleyard, D. (1970). Styles and methods of structuring a city. *Environment and behavior*, *2*(1), 100-117.

Allen, G. L., & Golledge, R. G. (2007). Wayfinding in urban environments. Threats from Car Traffic to the Quality of Urban Life: Problems, Causes, and Solutions, 79-102.

Atkinson, R.C., & Shiffrin, R.M. (1968). Human memory: A proposed system and its control processes. *The psychology of learning and motivation, 2*, 89-195.

Ballinger, G. A. (2004). Using generalized estimating equations for longitudinal data analysis. *Organizational research methods* , *7*(2), 127-150.

Barabási, A. L. (2016). *Network science*. Cambridge university press.

Bell, M. G., & Iida, Y. (1997). Transportation network analysis.

Berghauser-Pont, M. Y., & Haupt, P. (2010). *Spacematrix: space, density and urban form*. Rotterdam: NAi Publishers.

Blondel, V. D., Guillaume, J. L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, *2008*(10).

Bogers, E.A.I., M. Bierlaire, and S.P. Hoogendoorn, *Modeling learning in route choice*, in *Transportation Research Record*. 2007. p. 1-8.

Bovy, P. H. L., & Stern, E. (1990). Route choice: wayfinding in travel networks.

Brandes, U. (2001). A faster algorithm for betweenness centrality, The Journal of Mathematical Sociology, 25:2, 163-177.

Broach, J., Dill, J., & Gliebe, J. (2012). Where do cyclists ride? A route choice model developed with revealed preference GPS data. *Transportation Research Part A: Policy and Practice*, *46*(10), 1730-1740.

Buning, R. J., & Gibson, H. J. (2016). The role of travel conditions in cycling tourism: Implications for destination and event management. *Journal of Sport & Tourism*, *20*(3-4), 175-193.

Caldeira, A. M., & Kastenholz, E. (2019). Spatiotemporal tourist behaviour in urban destinations: a framework of analysis. *Tourism Geographies*, 1-29.

CBS (2019) Personenmobiliteit in Nederland; reiskenmerken en vervoerwijzen, regio's https://opendata.cbs.nl/statline/#/CBS/nl/dataset/83498ned/table?ts=1553677009195 (visited 2019-03-27).

Cenani, S., Arentze, T. A., & Timmermans, H. J. (2012). A cognitive learning model for dynamic activity-travel patterns. *Procedia-Social and Behavioral Sciences* , *5 4*, 580-588.

Cenani, S., T.A. Arentze, and H.J.P. Timmermans, Agent-based modeling of cognitive learning of dynamic activity-travel patterns. Transportation Letters, 2013. 5(4): p. 184-199.

Cenani, S., *Modeling Cognitive Learning of Urban Networks in Daily Activity-Travel Behavior*. 2013, Eindhoven University of Technology: Eindhoven.

Chamizo, V. D. (2003). Acquisition of knowledge about spatial location: Assessing the generality of the mechanism of learning. *The Quarterly Journal of Experimental Psychology: Section B* , *5 6*(1), 102-113.

Chrastil, E. R., & Warren, W. H. (2013). Active and passive spatial learning in human navigation: acquisition of survey knowledge. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *39*(5), 1520–37.

Cohen, I., Sebe, N., Garg, A., Chen, L. S., & Huang, T. S. (2003). Facial expression recognition from video sequences: temporal and static modeling. Computer Vision and image understanding, 91(1-2), 160-187.

Conzen, M. R. G. (1960). Alnwick, Northumberland: a study in town-plan analysis. *Transactions and Papers (Institute of British Geographers)*, (27), 111-122.

Couclelis, H., Golledge, R. G., Gale, N., & Tobler, W. (1987). Exploring the anchor-point hypothesis of spatial cognition. *Journal of Environ- mental Psychology*, *7*(2), 99-122.

Dickinson, J. E., Lumsdon, L. M., & Robbins, D. (2011). Slow travel: Issues for tourism and climate change. *Journal of Sustainable Tourism*, *19*(3), 281-300.

Dijst, M., & Vidakovic, V. (2000). Travel time ratio: the key factor of spatial reach. *Transportation*, *27*(2), 179-199.

Dogu, U., & Erkip, F. (2000). Spatial factors affecting wayfinding and orientation: A case study in a shopping mall. Environment and behavior, 32(6), 731-755.

Downs, R. M. & Stea, D. 1973 The Development of Spatial Cognition. In Image and Environment: Cognitive Mapping and Spatial Behavior. Edward Arnold. P221-225.

Duany, A., & Talen, E. (2002). Transect planning. *Journal of the Ameri- can Planning Association*, *68*(3), 245-266.

Duckham, M., Winter, S., & Robinson, M. (2010). Including landmarks in routing instructions. *Journal of Location Based Services*, *4*(1), 28-52.

East, D., Osborne, P., Kemp, S., & Woodfine, T. (2017). Combining GPS & survey data improves understanding of visitor behaviour. *Tourism Management*, *61*, 307-320.

Ekman, P., Friesen, W. V., & Ellsworth, P. (2013). Emotion in the human face: Guidelines for research and an integration of findings (Vol. 11). Elsevier.

Faulks, P., Ritchie, B. W., & Fluker, M. (2006). Cycle tourism in Australia: an investigation into its size and scope. Cooperative Research Centre for Sustainable Tourism.

Fedorova, T., De Graaff, L. & Sleutjes, B. (2018). Toerisme MRA 2017-2018. Gemeente Amsterdam, Onderzoek, Informatie en Statistiek.

Fietsersbond (2013) https://www.fietsersbond.nl/nieuws/ fietsende-toeristen-als-plaag/ (retrieved May 13 2019).

Gärling, T., Böök, A., & Ergezen, N. (1982). Memory for the spatial layout of the everyday physical environment: Differential rates of acquisition of different types of information. *Scandinavian Journal of Psychology* , *2 3*(1), 23-35.

Goldin, S. E., & Thorndyke, P. W. (1982). Simulating navigation for spatial knowledge acquisition. *Human factors*, *24* (4), 457-471.

Golledge, R. G., Dougherty, V., & Bell, S. (1995). Acquiring Spatial Knowledge: Survey Versus Route-Based Knowledge in Unfamiliar Environments. *Annals of the Association of American Geographers*, *85*(1), 134–158.

Golledge, R.G. and R.J. Stimson, Spatial behaviour: a geographic perspective. Spatial behaviour: a geographic perspective, 1997.

Golledge, R.G., Wayfinding Behavior: Cognitive Mapping and Other Spatial Processes, 1999.

Gong, X. (2016). Exploring Human Activity Patterns Across Cities through Social Media Data.

Gonzalez, M. C., Hidalgo, C. A., & Barabasi, A. L. (2008). Understanding individual human mobility patterns. *nature*, *453*(7196), 779.

Graham, C. (2008). Happiness and health: Lessons—and questions—for public policy. *Health affairs*, *27*(1), 72-87.

Haggett, P., & Chorley, R. J. (1969). *Network analysis in geography* (Vol. 1). Hodder Education.

Han, H., Meng, B., & Kim, W. (2017). Emerging bicycle tourism and the theory of planned behaviour. *Journal of Sustainable Tourism*, *25*(2), 292-309.

Hancock, T. O., Hess, S., & Choudhury, C. F. (2018). Decision field theory: Improvements to current methodology and comparisons with standard choice modelling techniques. *Transportation Research Part B: Methodological*, *107*, 18-40.

Hanley, J. A., Negassa, A., Edwardes, M. D. D., & Forrester, J. E. (2003). Statistical analysis of correlated data using generalized estimating equations: an orientation. *American journal of epidemiology*, *1 57* (4), 364-375.

Hasan, S., Schneider, C. M., Ukkusuri, S. V., & González, M. C. (2013). Spatiotemporal patterns of urban human mobility. *Journal of Statistical Physics*, *151*(1-2), 304-318.

Hayllar, B., & Griffin, T. (2005). The precinct experience: a phenomenological approach. *Tourism Management*, *26*(4), 517-528.

Hillier, B., & Hanson, J. (1984). The social logic of space, 1984. *Cam- bridge: Press syndicate of the University of Cambridge*.

Hirtle, S. C., & Hudson, J. (1991). Acquisition of spatial knowledge for routes. *Journal of environmental psychology*, *1 1* (4), 335-345.

Hartley, T., Maguire, E. A., Spiers, H. J., & Burgess, N. (2003). The well-worn route and the path less traveled: distinct neural bases of route following and wayfinding in humans. *Neuron*, *37*(5), 877-888.

Hoogendoorn-Lanser, S., & Van Nes, R. (2004). Multimodal choice set composition: Analysis of reported and generated choice sets. *Transportation research record*, *1898*(1), 79-86.

Isaacson, M., & Shoval, N. (2006). Application of tracking technologies to the study of pedestrian spatial behaviour. *The Professional Geographer*, *58*(2), 172-183.

Ishikawa, T., & Montello, D. R. (2006). Spatial knowledge acquisition from direct experience in the environment: Individual differences in the development of metric knowledge and the integration of separately learned places . *Cognitive psychology*, *52* (2), 93-129.

Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern recognition letters*, *31*(8), 651-666.

Kádár, B., & Gede, M. (2013). Where do tourists go? Visualizing and analysing the spatial distribution of geotagged photography. *Cartographica: The International Journal for Geographic Information and Geovisualization*, *48*(2), 78-88.

Kaplan, S., Manca, F., Nielsen, T. A. S., & Prato, C. G. (2015). Intentions to use bike-sharing for holiday cycling: An application of the Theory of Planned Behaviour. *Tourism Management*, *47*, 34-46.

Kaufman, L., & Rousseeuw, P. J. (2009). *Finding groups in data: an introduction to cluster analysis* (Vol. 344). John Wiley & Sons.

Kelly, J. W., & McNamara, T. P. (2010). Reference frames during the acquisition and development of spatial memories. *Cognition*, *116* (3), 409-420.

Kettunen, P., Irvankoski, K., Krause, C. M., Sarjakoski, T., & Sarjakoski, L. T. (2012). Geospatial images in the acquisition of spatial knowledge for wayfinding. *Journal of Spatial Information Science*, *2 012* (5), 75-106.

Kiefer, P., Giannopoulos, I., & Raubal, M. (2014). Where am I? Investigating map matching during self-localization with mobile eye tracking in an urban environment. *Transactions in GIS*, *18*(5), 660-686.

Kirasic, K. C. (2018). Ageing and spatial behaviour in the elderly adult. *Cognitive Mapping: Past, present and future.*

Klasnja, P., Consolvo, S., Choudhury, T., Beckwith, R., & Hightower, J. (2009, May). Exploring privacy concerns about personal sensing. *In International Conference on Pervasive Computing* (pp. 176-183). Springer, Berlin, Heidelberg.

Klippel, A. (2003, September). Wayfinding choremes. In *International Conference on Spatial Information Theory* (pp. 301-315). Springer Berlin Heidelberg.

Kuipers, B. (1978). Modeling spatial knowledge. *Cognitive science*, *2* (2), 129-153.

Lamont, M., & Causley, K. (2010). Guiding the Way: Exploring cycle tourists' needs and preferences for cycling route maps and signage. *Annals of Leisure Research*, *13*(3), 497-522.

Lang, J.T., Creating architectural theory: the role of the behavioral sciences in environmental design,. 1987, New York, Van Nostrand Reinhold Co.

Larsen, J., Axhausen, K. W., & Urry, J. (2006). Geographies of social networks: meetings, travel and communications. *Mobilities*, *1*(2), 261-283.

Lee, C. F. (2014). An investigation of factors determining cycling experience and frequency. *Tourism Geographies*, *16*(5), 844-862.

Lee, H., Cho, A., Jo, Y., & Whang, M. (2017). The Relationships Between Behavioral Patterns and Emotions in Daily Life. *In Advances in Computer Science and Ubiquitous Computing* (pp. 1332-1339). Springer, Singapore.

Liang, K. Y., & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, *73* (1), 13-22.

Lorenzo, G. D., Reades, J., Calabrese, F., & Ratti, C. (2012). Predicting personal mobility with individual and group travel histories. *Environment and Planning B: Planning and Design*, *39*(5), 838-857.

Lumsdon, L. (2000). Transport and tourism: cycle tourism–a model for sustainable development? *Journal of Sustainable Tourism*, *8*(5), 361-377.

Lynch, K., The Image of a City. 1960.

MacEachren, A. M. (1992). Application of environmental learning theory to spatial knowledge acquisition from maps. *Annals of the Association of American Geographers*, *82* (2), 245-274.

Manaugh, K., & El-Geneidy, A. M. (2012). What makes travel 'local' Defining and understanding local travel behavior. Journal of Transport and Land Use, 5(3), 15-27.

Manski, C. F. (1977). The structure of random utility models. *Theory and decision*, *8*(3), 229.

Marshall, S. (2004). *Streets and patterns*. Routledge.

Miller, H. J. (2005). A measurement theory for time geography. *Geographical analysis* , *37* (1), 17-45.

Mckercher, B., & Lau, G. (2008). Movement patterns of tourists within a destination. *Tourism geographies*, *10*(3), 355-374.

Meijles, E. W., De Bakker, M., Groote, P. D., & Barske, R. (2014). Analysing hiker movement patterns using GPS data: Implications for park management. *Computers, Environment and Urban Systems*, *47*, 44-57.

Menghini, G., Carrasco, N., Schüssler, N., & Axhausen, K. W. (2010). Route choice of cyclists in Zurich. *Transportation research part A: policy and practice*, *44*(9), 754-765.

Miller, H. J. (2005). A measurement theory for time geography. *Geographical analysis*, *37*(1), 17-45.

Modsching, M., Kramer, R., Hagen, K. T., & Gretzel, U. (2008). Using location-based tracking data to analyze the movements of city tourists. *Information Technology & Tourism*, *10*(1), 31-42.

Montello, D. R. (1997). The perception and cognition of environmental distance: Direct sources of information. In *International Conference on Spatial Information Theory* (pp. 297-311). Springer Berlin.

Montello, D. R. (1998). A new framework for understanding the acquisition of spatial knowledge in large-scale environments. *Spatial and temporal reasoning in geographic information systems*, 143-154.

Morello, E., & Ratti, C. (2009). A digital image of the city: 3D isovists in Lynch's urban analysis. *Environment and Planning B: Planning and Design*, *36*(5), 837-853.

NBTC Holland Marketing (2019) Perspectief bestemming Nederland 2030.

Newman, M. E. J. (2006). "Modularity and community structure in networks". *Proceedings of the National Academy of Sciences of the United States of America*. **103** (23): 8577–8696.

Newman, M. (2010). *Networks: an introduction*. Oxford university press.

Newsome, T. H., Walcott, W. A., & Smith, P. D. (1998). Urban activity spaces: Illustrations and application of a conceptual model for integrating the time and spacedimensions. *Transportation*, *25*(4),357-377.

NOS (2015) https://nos.nl/artikel/2073698-amsterdam-wil -toerist-door-fietsjungle-loodsen-met-komisch-filmpje.html (retrieved May 13 2019).

Oliveira, V. (2013). Morpho, a methodology for assessing urban form. *Urban Morpohology* 17:149-161.

Oliveira, V. and Medeiros, V. (2016) Morpho: Combining Morphological Measures. *Environment and Planning B: Planning and Design.*

Passini, R. (1981). Wayfinding: A conceptual framework. Urban Ecology, 5(1), 17-31.

Pettersson, R., & Zillinger, M. (2011). Time and space in event behaviour: Tracking visitors by GPS. *Tourism Geographies*, *13*(1), 1-20.

Rasmussen, T. K., Ingvardson, J. B., Halldórsdóttir, K., & Nielsen, O. A. (2015). Improved methods to deduct trip legs and mode from travel surveys using wearable GPS devices: A case study from the Greater Copenhagen area. *Computers, Environment and Urban Systems*, *54*, 301-313.

Raun, J., Ahas, R., & Tiru, M. (2016). Measuring tourism destinations using mobile tracking data. *Tourism Management*, *57*, 202-212.

Renso, C., Baglioni, M., de Macedo, J. A. F., Trasarti, R., & Wachowicz, M. (2013). How you move reveals who you are: understanding human behaviour by analyzing trajectory data. *Knowledge and information systems*, *37*(2), 331-362.

Richter, K. F. (2007). A uniform handling of different landmark types in route directions. In *International Conference on Spatial Information Theory* (pp. 373-389). Springer Berlin Heidelberg.

Richter, K. F.; Winter, S. (2014): Landmarks: GIScience for Intelligent Services. Springer, Cham, Switzerland.

Ritchie, B. W., & Hall, C. M. (1999). Bicycle tourism and regional development: A New Zealand case study. *Anatolia*, *10*(2), 89-112.

Ritchie, B. W., Tkaczynski, A., & Faulks, P. (2010). Understanding the motivation and travel behaviour of cycle tourists using involvement profiles. *Journal of Travel & Tourism Marketing*, *27*(4), 409-425.

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.

Ruddle, R. A., Volkova, E., & Bülthoff, H. H. (2011). Walking improves your cognitive map in environments that are large-scale and large in extent. *ACM Transactions on Computer-Human Interaction*, 18(2).

Sadalla, E. K., Burroughs, W. J., and Staplin, L. J. (1980) Reference points in spatial cognition. Journal of Experimental Psychology: Human Learning and Memory 6:516-528.

Schneider, F., Ton, D., Zomer, L. B., Hoogendoorn-Lanser, S., Duives, D., Cats, O., & Hoogendoorn, S. (2020). Trip Chain Complexity: A Comparison among Latent Classes of Daily Mobility Patterns. *Transportation.*

Schönfelder, S., & Axhausen, K. W. (2016). *Urban rhythms and travel behaviour: spatial and temporal phenomena of daily travel*. Routledge.

Schwanen, T., & Dijst, M. (2002). Travel-time ratios for visits to the workplace: the relationship between commuting time and work duration. *Transportation Research Part A: Policy and Practice*, *36*(7), 573-592.

Selkirk, K. E. (1982). *Pattern and place: an introduction to the mathematics of geography*. Cambridge University Press.

Sevtsuk, A., & Mekonnen, M. (2012). Urban network analysis. *Revue internationale de géomatique*, 287-305.

Shoval, N., & Isaacson, M. (2007). Tracking tourists in the digital age. *Annals of Tourism Research*, *34*(1), 141-159.

Shoval, N., & Isaacson, M. (2009). *Tourist mobility and advanced tracking technologies*. Routledge.

Shoval, N., McKercher, B., Ng, E., & Birenboim, A. (2011). Hotel location and tourist activity in cities. *Annals of Tourism Research*, *38*(4), 1594-1612.

Shoval, N., & Raveh, A. (2004). Categorization of tourist attractions and the modeling of tourist cities: based on the co-plot method of multivariate analysis. *Tourism Management*, *25*(6), 741-750.

Shoval, N., Schvimer, Y., & Tamir, M. (2018). Tracking technologies and urban analysis: adding the emotional dimension. *Cities*, *72*, 34-42.

Siegel, A.W. and S.H. White, *The Development of Spatial Representations of Large-Scale Environments.* Advances in Child Development and Behavior, 1975. **10**(C): p. 9-55.

Skov-Petersen, H., Barkow, B., Lundhede, T., & Jacobsen, J. B. (2018). How do cyclists make their way?-A GPS-based revealed preference study in Copenhagen. *International Journal of Geographical Information Science*, *32*(7), 1469-1484.

Stanney, K. M., Cohn, J., Milham, L., Hale, K., Darken, R., & Sullivan, J. (2013). Deriving training strategies for spatial knowledge acquisition from behavioral, cognitive, and neural foundations. *Military Psychology* , *25* (3), 191-205.

Song, C., Qu, Z., Blumm, N., & Barabási, A. L. (2010). Limits of predictability in human mobility. *Science*, *327*(5968), 1018-1021.

Srinivas, S. (2010). *Influence of motivation on wayfinding* (Doctoral dissertation, University of Pittsburgh).

Stern, E., & Leiser, D. (1988). Levels of spatial knowledge and urban travel modeling. Geographical Analysis, 20 (2), 140-155.

Stevens, Q.. "*The Shape of Urban Experience: a Re-Evaluation of Lynch's Five Elements*." In *Evaluation in Progress-Strategies for Environmental Research and Implementation*. IAPS. Vienna, Austria: Österreichischer Kunst- und Kulturverlag, 2004.

Susilo, Y. O., & Dijst, M. (2009). How Far Is Too Far? Travel Time Ratios for Activity Participation in the Netherlands. *Transportation Research Record*, *2134*(1), 89-98.

Ton, D., Cats, O., Duives, D., & Hoogendoorn, S. (2017). How Do People Cycle in Amsterdam, Netherlands?: Estimating Cyclists' Route Choice Determinants with GPS Data from an Urban Area. *Transportation research record*, *2662*(1), 75-82.

Ton, D., Duives, D. C., Cats, O., Hoogendoorn-Lanser, S., & Hoogendoorn, S. P. (2019). Cycling or walking? Determinants of mode choice in the Netherlands. *Transportation Research Part A: Policy and Practice*, *123*, 7-23.

Ton, D., Zomer, L. B., Schneider, F., Hoogendoorn-Lanser, S., Duives, D., Cats, O., & Hoogendoorn, S. (2019). Latent classes of daily mobility patterns: the relationship with attitudes towards modes. *Transportation*, 1-24.

Ton, D. (2019). *Unravelling Mode and Route Choice Behaviour of Active Mode Users* (Doctoral dissertation, Delft University of Technology).

Van Nes, A., Berghauser Pont, M., & Mashhoodi, B. (2012). Combination of Space syntax with spacematrix and the mixed use index: The Rotterdam South test case. In *8th International Space Syntax Symposium, Santiago de Chile, Jan. 3-6, 2012*. PUC, Santiago, Chili.

van der Spek, S. (2010). Tracking tourists in historic city centres. In *Information and Communication Technologies in Tourism 2010* (pp. 185-196). Springer, Vienna.

Van Wee, B. (2002). Land use and transport: research and policy challenges. *Journal of transport geography*, *10*(4), 259-271.

Vermunt, J. K., & Magidson, J. (2002). Latent class cluster analysis. Applied latent class analysis, 11, 89-106.

Vilar, E., & Rebelo, F. (2008). Virtual reality in wayfinding studies. *Virtual Reality*.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *nature*, *393*(6684), 440.

Willis, K. S., Hölscher, C., Wilbertz, G., & Li, C. (2009). A comparison of spatial knowledge acquisition with maps and mobile maps. *Computers, Environment and Urban Systems*, *3 3*(2), 100-110.

Woollett, K., & Maguire, E. A. (2011). Acquiring "the Knowledge" of London's layout drives structural brain changes. *Current biology* ,*21*(24),2109-2114.

World Tourism Organization (2018a), UNWTO Tourism Highlights, 2018 Edition, UNWTO, Madrid.

World Tourism Organization (UNWTO); Centre of Expertise Leisure, Tourism & Hospitality; NHTV Breda University of Applied Sciences; and NHL Stenden University of Applied Sciences (eds., 2018b), *'Overtourism'? – Understanding and Managing Urban Tourism Growth beyond Perceptions*, UNWTO, Madrid.

Xia, J. C., Zeephongsekul, P., & Arrowsmith, C. (2009). Modelling spatio-temporal movement of tourists using finite Markov chains. *Mathematics and Computers in Simulation*, *79*(5), 1544-1553.

Zacharias, J. (1997). The impact of layout and visual stimuli on the itineraries and perceptions of pedestrians in a public market. *Environment and Planning B: Planning and Design*, *24*(1), 23-35.

Zheng, W., Huang, X., & Li, Y. (2017). Understanding the tourist mobility using GPS: Where is the next place?. *Tourism Management*, *59*, 267-280.

Zhong, C., Manley, E., Arisona, S. M., Batty, M., & Schmitt, G. (2015). Measuring variability of mobility patterns from multiday smart-card data. *Journal of Computational Science*, 9, 125-130.

Zillinger, M. (2007). Tourist routes: A time-geographical approach on German car-tourists in Sweden. *Tourism Geographies*, *9*(1), 64-83.

Zomer, L. B., Daamen, W., Meijer, S., & Hoogendoorn, S. P. (2015). Managing crowds: The possibilities and limitations of crowd information during urban mass events. In *Planning Support Systems and Smart Cities* (pp. 77-97). Springer, Cham.

Summary

Introduction

Every day residents and visitors find their way through complex urban networks to go to work, reach school, or go sightseeing. The density of the urban street fabric poses challenges to travellers compared to the more sparse national highway or public transport system. While it is rare to get lost on your daily commute, it is quite common to deviate from the shortest or fastest route, or to use navigation to avoid congestion and to conveniently find your way.

In the Netherlands active modes (pedestrians and cyclists) account for more than 50% of all urban trips. Active modes have been promoted as a sustainable, healthy and inexpensive means of transport that could mitigate urban congestion and urban livability issues due to increased urbanization. Therefore, urban planners and policy makers are looking for ways to create walkable and bikeable cities. To provide people with understandable urban networks and travel information, the complexity of human behaviour requires a deeper understanding of how people find the way by foot and bicycle and identification of the role of spatial (network) knowledge. Wayfinding behaviour in cities is defined by the strategies that people employ to decide (how) to move from one place to another within an urban area.

The purpose of the dissertation study is to narrow the gap between research on travel behaviour research and research on urban spatial knowledge. Within transport science it is common to estimate and predict travel behaviour using discrete choice or activity-travel models, because of well-defined descriptive and data collection procedures. These methods assume, to a large extent, that decision-making behaviour in travel behaviour is hierarchical and sequential. We hypothesize that due to this lack of knowledge existing route choice paradigms are behaviourally inadequate to model mobility choices of tourists and newcomers, as they rely more on the generation of a choice set. Regarding the understanding of urban wayfinding, a theory is still absent which can describe when, and why, a particular route or

wayfinding landmark is part of a (network) choice set in relation to the (learning of) urban environment, mobility patterns, and information acquisition behaviour of (active) travellers. This dyad will for example become problematic in the future as predictions, based on economic prosperity and cheap long-haul travel costs, estimate a growth of 44% to 200%, yielding 28.8 to 41.9 million, tourists in The Netherlands by 2030. As currently 40% (8 million) of the tourists stay in the capital city Amsterdam this becomes what is called "overtourism" when the unequal dispersion of tourists remains. To advance the understanding of bicycling behaviour of tourists, thorough insights are required into activity and movement patterns of tourists and how choices and patterns evolve over time.

Each trip requires people to make various decisions before and during travelling. These decisions pertain to the modes and routes to be used, and which activities will be performed where and when. Due to individual differences in navigational preferences (minimize turns and thus choosing a simpler yet longer route) and socio-demographic characteristics (gender, age, and mode availability) the urban experience differs, and as a consequence, the mental representation of the environment (perceived accessibility levels, and salient areas) is likely to be different. In turn, these differences will influence the amount of exploration or habitual travelling during future trips. All these characteristics evolve around the wayfinding attitude or style, defined by the strategies that people use to decide how to move from one place to another (Montello 1995). It relates to the set of preferences, selection, and application of navigational strategies, the attitude towards travelling, and the ability to reach the intended destination. As such, differences in travel behaviour are expected to determine the extent to which wayfinding styles and navigational preferences are important to individuals. This dissertation focuses on spatial knowledge development during the exploration of a city.

Urban wayfinding styles

Differences in urban wayfinding behaviour in relation to daily travel patterns are important to understand route choice behaviour, identify potential navigation problems, design more legible cities (understandable, imagable and coherent urban environments), and provide comprehensible travel information. Therefore, the goal of this chapter is to jointly investigate the differences between urban wayfinding styles and the relations with socio-demographic, motility, urban environment, navigational preferences, and daily travel behaviour.

The findings of this chapter are based on a sample of the Dutch population of 1101 respondents. All respondents completed a three-day travel diary as part of the Mobility Panel Netherlands, and an additional cross-sectional survey designed to capture perceptions, attitudes, and wayfinding for active modes. A Factor Analysis is conducted to identify urban wayfinding styles based on a Dutch version of the self-report questionnaire of environmental spatial skills originally developed in Santa Barbara (SBSOD). Generalized Linear Models are used to estimate to what extent various determinants affect two hypothesized urban wayfinding styles, in this chapter coined as Orientation Ability and Knowledge Gathering & Processing Ability.

The main findings of the study are an associated effect of gender and age on both urban wayfinding styles. On average we observe both higher Orientation Ability as Knowledge Gathering and Processing Ability among men and with increased age. While the navigational preference to follow the bearing line (direct straight distance between origin and destination) and average daily distance travelled by car have disassociated effects. The remaining determinants are only significant in either Orientation Ability or Knowledge Gathering & Processing Ability, providing evidence that mainly different processes describe each wayfinding style.

Urban salient areas

Spatial urban route knowledge consists of the internalized representation of a sequence of actions to be performed at certain locations, cued by wayfinding landmarks. Determining the location of distinctive landmarks is thus important in research on route choice, urban cognition, and travel information. Currently, most approaches to identify landmarks require vast data collection efforts. To overcome these demands, this chapter proposes a spatial analytic method able to leverage on open-source datasets to identify urban wayfinding landmarks as salient urban areas.

The method consists of five steps based on data management, grouping analysis, and cluster and outlier analysis. Determinants to identify salient urban areas are building volume, surface, height, building year, and the number of buildings in a 100 square meters grid-cell.

Findings have been applied to identify differences in distribution of clustering and dispersion between local and global salient urban areas using the Gini coefficient, based on an open-source GIS dataset on the built environment of Amsterdam.

Urban activity patterns of tourists

Until 2030 the expected urban tourism growth may lead to "overtourism" in city centres, with excessive noise, nuisance for inhabitants, and pressure on infrastructure. Therefore, global strategies and measures to better understand and manage urban tourism have been established previously. However, the effect of the measures heavily depends on the travel behaviour of tourists within the respective cities. More insights are necessary to understand how urban travel behaviour of tourists can be influenced to i) stimulate new itineraries (activity sequences), ii) promote spatial dispersal, and iii) promote temporal dispersal.

To this end, we develop metrics and methods that can be used to characterise the spatiotemporal travel and activity patterns of tourists with access to a bicycle in metropolitan areas. The analyses are based on GPS data of tourists' cycling behaviour collected in the Amsterdam Metropolitan Region in July and August 2017. Empirical insights are provided based on 10.342 processed activities. The study entails a k-means clustering to identify 105 activity zones, network analyses to investigate the existence and behaviour of activity communities, observed activity spaces are related with travel time ratio to identify differences in activity patterns, and temporal activity patterns shed light on possibilities for distribution of tourists over the day.

The findings indicate that i) consistency between the activity network of tourists with access to a bicycle and expectations from other transport networks, and ii) mobility patterns of tourists using bicycles in metropolitan regions varies from activity oriented to trip orientated, and from directed to compact. The insights to improve urban design and travel information for tourists can be used by policy-makers and urban planners to nudge tourists to less crowded areas. In particular, possibilities are discussed to achieve wider spatial and temporal dispersal of tourists if they travel by bicycle.

Urban routing patterns and the influence of spatial learning

This chapter presents the evolution of spatial behaviour when new knowledge is acquired with successive bicycle trips of tourists. This is done by modelling the influence of experienced travel behaviour and spatial knowledge acquisition on the route selection space of tourists travelling by bicycle in Amsterdam. To be able to support policy makers to gain more insights into the impacts of these tourist flows, a greater understanding of their urban travel behaviour is needed. Generalized Estimating Equations are used to assess the (spatial) learning effect of tourists as a function of the approximated trip purpose, familiarity, and movement patterns. Four route selection space characteristics are investigated in this research: detour ratio, maximum deviation from the bearing line, eccentricity, and curvature.

The spatial density probabilities signify that tourism routing behaviour differs from habitual routing behaviour in Amsterdam. Regarding learning effects of tourists, the spatial density probabilities of the route selection space contracts with each trip as well as with increased Euclidean direct distance between origin and destination. We show that the four route selection space models depend mainly on trip purpose (new or recurrent destination, travel time ratio of the trip, and the number of activities that will be visited in a day). The route selection space is likely to be inflated during the first few trips travelling to a new destination where the activity duration is lower compared to the travel time to get there. Due to learning effects the route selection space contracts when approaching the end of the activity sequence. For example, after the 5th or 6th trip, visiting a new destination does not correspond to a larger deviation or detour.

Moreover, to model route choices in cities with large tourists flows, the proposed route selection space metrics can be valuable in defining the spatial boundaries for generation of the set of feasible route alternatives.

Study contributions and implications

The overarching objective of the studies in this dissertation is to unravel how travel behaviour, urban environments, and information services relate to spatial knowledge development. The conclusions, in terms of contributions and implications, of are fourfold:

We tested a wayfinding theory based on existing literature to identify the components of urban wayfinding behaviour, and investigated the relation with daily travel behaviour based on travel diary data

The Factor Analysis reveals two components of urban wayfinding wayfinding, coined Orientation Ability (attitude and basic skills to be able to orient and navigate effectively in an urban environment) and Knowledge Gathering & Processing Ability (attitude and preferences to extend knowledge about the environment, e.g. explore cities and take new routes). The findings clarify the extent to which these wayfinding styles relate to daily travel behaviour (e.g. activity and route choice behaviour) in the urban environment.

From the literature background it is hypothesized that the total average travel distance (by car and foot) has a negative relation with the wayfinding score. The results in this chapter also show a negative relation for distance travelled by car, and for the first time, also distance cycled by men with Orientation Ability. On the contrary, this chapter also shows that the total average distance travelled by car and the interaction effect between average number of car trips and active navigation ratio have positive relations with Knowledge Gathering & Processing Ability. Although the majority of the research found in literature investigates pedestrian wayfinding, the distance travelled by foot and public transport are not significant in this chapter.

In relation to legible urban wayfinding, future research should investigate if indeed people with high Orientation Ability are more likely to correspond to higher variability in the streets of chosen routes. This implies that people with lower levels of Orientation Ability will compensate for the complexity of the urban wayfinding task by preferring a longer route along familiar streets. Thus, even if high connectivity exists, but all people have low orientation abilities, still not much route variation will occur and it will become more difficult to mitigate congestion and distribute large cyclists flows more evenly.

Both hypothesised wayfinding styles can be used complementary as different processes influence them. However, two determinants (navigational preference to follow the bearing line and average daily distance travelled by car) have an op effect on both wayfinding styles. This could indicate a trade-off, because gathering and processing more spatial knowledge will ultimately require more orientation ability in order to process the knowledge into usable wayfinding styles. The navigational preference to follow the bearing line is not beneficial when there is a low amount of spatial knowledge, as this does not encourage the acquisition of more spatial knowledge. If a satisfactory amount of spatial knowledge has been acquired using the bearing line as a navigational preference is useful to reduce the workload.

We developped a methodology based on spatial analytics to use open spatial data to characterize salient and legible areas in an urban environment.

The implications of identification of salient urban areas can provide new insights on how to analyze how wayfinding landmarks structure environmental knowledge and investigate influences on wayfinding strategies. This environmental knowledge (configuration of landmarks) is assumed to become available when also knowledge has been memorized about the general interrelationships between landmarks. If people use these wayfinding landmarks as part of the wayfinding strategy, this is expected to be observable in their route choice behavior. For example it could be more likely that people take a detour if more wayfinding landmarks are passed.

We gained insights into activity patterns of tourists based on a large empirical field study to collect GPS trajectories of bicycles. The insights are used to develop new theories to better understand and influence travel behavior of tourists by bicycle in crowded cities.

Insights pertain to the activity network of tourists with access to a bicycle, which appears to be consistent with expectations from other transport networks; we observe that tourists tend to visit certain clusters of activities (communities) leading to differences in activity patterns, while there are at the same time strong connections between communities with one or two activities. Secondly, the combination of activity spaces, travel ratios, and travel pattern structure built on top of existing theories of tourists' travel behaviour. Mobility patterns of tourists using bicycles in metropolitan regions vary from activity oriented to trip orientated, and from directed to compact. Further research is required to explore the difference between residents and tourists, and tourists travelling by foot and public transport.

By collecting GPS trajectory data of bicycles of tourists and combining these with spatial maps from open data, we have computed determinants that describe the (development of) spatial knowledge level of tourists. A model is used to estimate to what extent the detour ratio and deviation area of a bicycle trip can be predicted based on the theoretic acquired spatial knowledge.

The spatial density probabilities of the route selection space contracts with every trip and increased Euclidean direct distance between origin and destination. The route selection space characteristics depend mainly on trip purpose (new or recurrent destination, travel time ratio, and the number of activities that will be visited in a day). Contrary to previous findings in literature based on habitual car trips, longer Euclidean direct distances between origin and destination decrease the relative detour and maximum deviation, and increase the centricity. The acquired knowledge of tourists can be captured by the current number of trips and number of new activities that have been visited. Here, it should be noted that the trip number is highly correlated with experienced bicycle distance, covered area, retraced area, anchor points, and recurrent activities. The Generalized Estimating Equation models also indicate that most learning effects stabilizes after 8 or more trips around the median. Combined, these insights suggest that experienced travel distance, covered area, retrace area, number of anchor points, and number of recurrent activities in a daily travel pattern may influence the route selection space of habitual trips.

The results of this chapter indicate that with longer Euclidean direct distance to reach activities in the outer areas, the detour ratio and maximum deviation decrease and trips become more centric. Also if intermediate activities are promoted to cut a long trip in two segments, increased trip number, number of activities that will be performed in a day and the activity sequence also correspond to a decrease in detour ratio and maximum deviation and increase of centricity. Therefore, it can be concluded that tourists need to be able to follow a bicycle route along a more directed street network to reach outer areas. This does not necessarily entail more streets, as tailored route advice for tourists also provides an opportunity, for example if it takes unique sites into account and allows for a larger detour from the shortest path.

About the Author

Lara-Britt Zomer was born in Leiden, The Netherlands on the Sunday of April 24th 1988. After finishing her high school at Onze Lieve Vrouwe Lyceum in Breda she pursued an architectural degree at Delft Univeristy of Technology from 2006. She completed her Bachelor degree in 2010, and also spent half a year at École nationale supérieure d'architecture de Paris-La Villette.

In 2011 she embarked on a new journey, and started a multidisciplinary M.Sc. programme, Transportation, Infrastructure and Logistics at the three faculties of Delft Univerity of Technology: Civil Engineering, Systems Engineering, Policy Analysis and Management, and Mechanical Engineering. She obtained her M.Sc. degree in the spring of 2014 with a research on crowd management and the effectiveness of information to advise visitors of urban mass events to choose a less crowded activity location.

In October 2015 she came back to Delft to start a Ph.D. at the department of Transport & Planning that is part of a larger European research project called ALLEGRO. Since November 2019 she is working at the Provincie Zuid-Holland, where she is, amongst others, the project leader public transport data monitoring.

On the last day of the Ph.D in Delft she initiated together with Paul van Gent and Giulia Reggiani a start-up called HappyStreets, to capture not only how many, but also how happy cyclists are, using cameras, smile detection & Artificial Intelligence models. HappyStreets gives you the insights and advise to make your streets happier!

Publications

Journal articles

- **Zomer, L.B.**, Daamen, W., Meijer, S., and Hoogendoorn, S. P. (2015). Managing crowds: The possibilities and limitations of crowd information during urban mass events. In *Planning Support Systems and Smart Cities* (pp. 77-97). Springer, Cham.
- Ton, D., **Zomer, L.B.**, Schneider, F., Hoogendoorn-Lanser, S., Duives, D., Cats, O., and Hoogendoorn, S. (2019). Latent classes of daily mobility patterns: the relationship with attitudes towards modes. *Transportation,* 1-24.
- **Zomer, L.B.**, Schneider, F., Ton, D., Hoogendoorn-Lanser, S., Duives, D., Cats, O., and Hoogendoorn, S. (2019). Determinants of urban wayfinding styles. *Travel Behaviour and Society*, *17*, 72-85.
- Schneider, F., Ton, D., **Zomer, L.B.**, Daamen, W., Duives, D., Hoogendoorn-Lanser, S., and Hoogendoorn, S. (2020). Trip chain complexity: a comparison among latent classes of daily mobility patterns. *Transportation*, 1-23.

The following articles are currently under review:

- **Zomer, L.B.**, Duives, D.C., Cats, O., and Hoogendoorn, S.P. (under review). Activity Patterns of Tourists in Amsterdam ft GPS Bicycle Data. *Travel Behaviour & Society.*
- **Zomer, L.B.**, Duives, D.C., Cats, O., and Hoogendoorn, S.P. (under review). On the Relation between Learning the City and Routing. *Transportation Research Part F: Travel behaviour and Psychology.*

Peer-reviewed conference contributions

- **Zomer, L. B.**, Moustaid, E., and Meijer, S. (2015, December). A meta-model for including social behavior and data into smart city management simulations. In *2015 Winter Simulation Conference (WSC)* (pp. 1705-1716). IEEE.
- **Zomer, L.B.**, Duives, D. C., Cats, O., and Hoogendoorn, S. P. (2017). Spatial Analytics for the Identification of Salient Urban Areas. In *15th International Conference on Computers in Urban Planning and Urban Management*. August 2017, Adelaide, Australia.
- Schneider, F., Ton, D., **Zomer, L.B.**, Daamen, W., Duives, D.C., Hoogendoorn-Lanser, S. and Hoogendoorn, S.P. (2018). Trip chains: a comparison among latent mobility pattern classes. *Presented at the International Association of Travel Behaviour Research Conference*, July 2018, Santa Barbara, USA.
- Ton, D., **Zomer, L.B.**, Schneider, F., Hoogendoorn-Lanser, S., Duives, D., Cats, O., and Hoogendoorn, S. (2018). Latent classes of daily mobility patterns: The relationship with attitudes towards modes – *Presented at the International Association of Travel Behaviour Research Conference*, July 2018, Santa Barbara, USA.
- **Zomer, L.B.**, Schneider, F., Ton, D., Hoogendoorn-Lanser, S., Duives, D., Cats, O., and Hoogendoorn, S. (2018). Wayfinding styles: The relationship with mobility patterns & navigational preferences – *Presented at the International Association of Travel Behaviour Research Conference*, July 2018, Santa Barbara, USA.
- **Zomer, L.B.**, Duives, D. C., Cats, O., and Hoogendoorn, S. P. (2019). On the relation between learning the city and routing. *Presented at the International Choice Modelling Conference*, August 2019, Kobe, Japan.
- Ton, D., **Zomer, L.B.**, Duives, D.C, and Hoogendoorn, S.P. (2020). GPS data and travel diaries: Two sides of one coin or two currencies? *To be presented at International Steering Committee for Transport Survey Conferences*, June 2020, Porto Novo Portugal.

Other articles, presentations and symposia

- **Zomer, L.B.**, and Langbroek, J. (2015). A Decade of Dutch Congestion Hit Parades: Data Analysis of an Integral Highway Extension Plan on Congestion Occurrences from 2002 to 2014. *Presented at the Swedish Transportation Research Conference*, October 2015, Karlstad, Sweden.

- **Zomer, L.B.**, Cats, O., and Hoogendoorn, S. P. (2016). Unravelling Urban Cognition: Quantification of Distortion of Perceived relative Locations. *Presented at the Symposium of the European Association for Research in Transportation*, September 2016, Delft, The Netherlands.
- **Zomer, L.B.**, Cats, O., and Hoogendoorn, S.P. (2016). Simulation approach for generating landmarks based on cognitive wayfinding strategies in urban environments. *Presented at the TRAIL conference*, November 2016, Utrecht, The Netherlands.
- **Zomer, L.B.**, Duives, D., Cats, O., and Hoogendoorn, S.P. (2017). The art of GPS bicycle data and activity identification. *Presented at the TRAIL conference*, November 2017, Utrecht, The Netherlands.
- Hoogendoorn, S.P., Ton, D., **Zomer, L.B.**, de Graaf, S., and Barmentlo, H. (2015). *Resurgence in the simulation of slow traffic.* NM Magazine (#3).
- **Zomer, L.B.**, Ton, D., and van Ooijen, T. (2018). *Active Modes First*. NM Magazine (#1).
- **Zomer, L.B.**, Duives, D., Cats, O., and Hoogendoorn, S.P. (2019) *Lessen uit Amsterdam: het managen van de fietstoerist.* NM Magazine (#2).
- **Zomer, L.B.**, Vial, A., Reggiani, G., Ton, D., Wierbos, M.J., Gong, V., Schneider, F., Feng, Y., Sparnaaij, M., Gavriilidou, A., van Oijen, T., Duives, D.C., Yuan, Y., Cats, O., Knoop, V.L., Daamen, W. and Hoogendoorn, S.P. (2019). The Impact of Cycling Research: Connecting Science and Practice. *Presented at Cycling Research Board*, October 2019, Delft, The Netherlands.

Data collection

- Perceived distances in Delft
- LUCY – Learning and Understanding of Cyclists behaviour (in collaboration with AMS and TSH-City)
- MPN special issue PAW-AM – Perceptions, Attitudes, and Wayfinding of Active Modes (in collaboration with KiM)
- TRACY – Tracking Cyclists during MATTS 2018 conference in Delft & CRB Conference 2018 in Amsterdam.

Trail Thesis Series

The following list contains the most recent dissertations in the TRAIL Thesis Series. For a complete overview of more than 275 titles see the TRAIL website: www.rsTRAIL.nl.

The TRAIL Thesis Series is a series of the Netherlands TRAIL Research School on transport, infrastructure and logistics.

Zomer, L.-B., *Unravelling Urban Wayfinding: Studies on the development of spatial knowledge, activity patterns, and route dynamics of cyclists,* T2021/16, May 2021, TRAIL Thesis Series, the Netherlands

Núñez Velasco, J.P., *Should I Stop or Should I Cross? Interactions between vulnerable road users and automated vehicles,* T2021/15, May 2021, TRAIL Thesis Series, the Netherlands

Duivenvoorden, K., *Speed Up to Safe Interactions: The effects of intersection design and road users' behaviour on the interaction between cyclists and car drivers,* T2021/14, April 2021, TRAIL Thesis Series, the Netherlands

Nagalur Subraveti, H.H.S., *Lane-Specific Traffic Flow Control,* T2021/13, March 2021, TRAIL Thesis Series, the Netherlands

Beirigo, B.A., *Dynamic Fleet Management for Autonomous Vehicles: Learning- and optimization-based strategies,* T2021/12, March 2021, TRAIL Thesis Series, the Netherlands

Zhang, B., *Taking Back the Wheel: Transition of control from automated cars and trucks to manual driving,* T2021/11, February 2021, TRAIL Thesis Series, the Netherlands

Boelhouwer, A., *Exploring, Developing and Evaluating In-Car HMI to Support Appropriate use of Automated Cars,* T2021/10, January 2021, TRAIL Thesis Series, the Netherlands

Li, X., *Development of an Integrity Analytical Model to Predict the Wet Collapse Pressure of Flexible Risers,* T2021/9, February 2021, TRAIL Thesis Series, the Netherlands

Li, Z., *Surface Crack Growth in Metallic Pipes Reinforced with Composite Repair System,* T2021/8, January 2021, TRAIL Thesis Series, the Netherlands

Gavriilidou, A., *Cyclists in Motion: From data collection to behavioural models,* T2021/7, February 2021, TRAIL Thesis Series, the Netherlands

Methorst, R., *Exploring the Pedestrians Realm: An overview of insights needed for developing a generative system approach to walkability,* T2021/6, February 2021, TRAIL Thesis Series, the Netherlands

Walker, F., *To Trust or Not to Trust? Assessment and calibration of driver trust in automated vehicles,* T2021/5, February 2021, TRAIL Thesis Series, the Netherlands

Schneider, F., *Spatial Activity-travel Patterns of Cyclists,* T2021/4, February 2021, TRAIL Thesis Series, the Netherlands

Madadi, B., *Design and Optimization of Road Networks for Automated Vehicles,* T2021/3, January 2021, TRAIL Thesis Series, the Netherlands

Krabbenborg, L.D.M., *Tradable Credits for Congestion Management: support/reject?,* T2021/2, January 2021, TRAIL Thesis Series, the Netherlands

Castelein, B., *Accommodating Cold Logistics Chains in Seaport Clusters: The development of the reefer container market and its implications for logistics and policy,* T2021/1, January 2021, TRAIL Thesis Series, the Netherlands

Huang, B., *The Influence of Positive Interventions on Cycling*, T2020/20, December 2020, TRAIL Thesis Series, the Netherlands

Xiao, L., *Cooperative Adaptive Cruise Control Vehicles on Highways: Modelling and Traffic Flow Characteristics,* T2020/19, December 2020, TRAIL Thesis Series, the Netherlands

Polinder, G.J., *New Models and Applications for Railway Timetabling,* T2020/18, December 2020, TRAIL Thesis Series, the Netherlands

Scharpff, J.C.D., *Collective Decision Making trough Self-regulation*, T2020/17, November 2020, TRAIL Thesis Series, the Netherlands

Guo, W. *Optimization of Synchromodal Matching Platforms under Uncertainties,* T2020/16, November 2020, TRAIL Thesis Series, the Netherlands

Narayan, J., *Design and Analysis of On-Demand Mobility Systems,* T2020/15, October 2020, TRAIL Thesis Series, the Netherlands

Gong, X., *Using Social Media to Characterise Crowds in City Events for Crowd Management,* T2020/14, September 2020, TRAIL Thesis Series, the Netherlands

Rijal, A., *Managing External Temporal Constraints in Manual Warehouses,* T2020/13, September 2020, TRAIL Thesis Series, the Netherlands

Alonso González, M.J., *Demand for Urban Pooled On-Demand Services: Attitudes, preferences and usage,* T2020/12, July 2020, TRAIL Thesis Series, the Netherlands

Alwosheel, A.S.A., *Trustworthy and Explainable Artificial Neural Networks for choice Behaviour Analysis,* T2020/11, July 2020, TRAIL Thesis Series, the Netherlands

Zeng, Q., *A New Composite Indicator of Company Performance Measurement from Economic and Environmental Perspectives for Motor Vehicle Manufacturers,* T2020/10, May 2020, TRAIL Thesis Series, the Netherlands

Mirzaei, M., *Advanced Storage and Retrieval Policies in Automated Warehouses,* T2020/9, April 2020, TRAIL Thesis Series, the Netherlands

Nordhoff, S., *User Acceptance of Automated Vehicles in Public Transport,* T2020/8, April 2020, TRAIL Thesis Series, the Netherlands