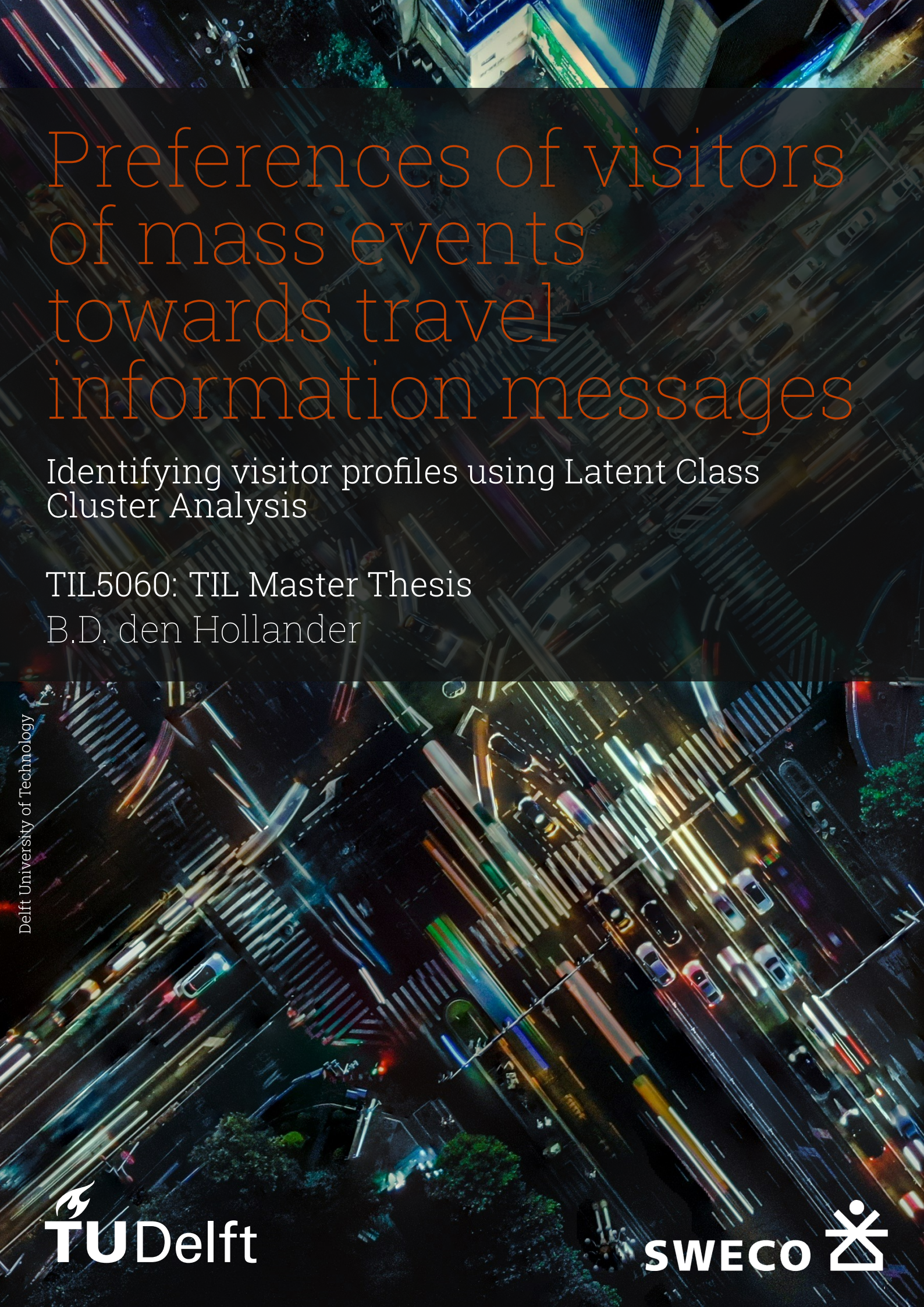

Preferences of visitors of mass events towards travel information messages

Identifying visitor profiles using Latent Class Cluster Analysis

TIL5060: TIL Master Thesis

B.D. den Hollander

Delft University of Technology

TU Delft

SWECO

Preferences of visitors of mass events towards travel information messages

Identifying visitor profiles using Latent Class Cluster Analysis

by

B.D. den Hollander

Student number - 4597494

to obtain the degree of Master of Science
at the Delft University of Technology,

Supervisors:

Dorine Duives	TU Delft
Maarten Kroesen	TU Delft
Serge Hoogendoorn	TU Delft
Merel Rozier	Sweco

Faculty:	Faculty of Civil Engineering and Geosciences, Delft
Project duration:	September 2022 - April 2023

TU Delft

SWECO

Preface

This document is my master thesis titled: *Preferences of visitors of mass events towards travel information messages - Identifying visitor profiles using Latent Class Cluster Analysis*. With this thesis, the Master Transport, Infrastructure and Logistics at Delft University of Technology will be completed.

My affinity with this topic already started during my bachelor Civil Engineering. I completed this bachelor's by writing a thesis about the walking behaviour of people in supermarkets during COVID-19. This research increased my interest in this topic and helped me choose a master's degree. While following this master's, I learned about the elective Active Modes. Even though I did not need the credits, I decided to take the course, where I learned more about people's sometimes unexpected travel choices. For the course Research Project, I chose to write a report about the influence of group formation on individual route choices at mass events. This made me realise that there is still much to learn about travel behaviour. During my master's thesis, I had the opportunity to contribute to this at Sweco. Sweco gave me the opportunity to work on the recently developed CROMAS. A project that intrigued me from day one.

I want to thank Sweco for the opportunity to perform this research at the company. The colleagues from the Mobility department always made me feel welcome and were happy to tell me more about the projects of this department. I would, in particular, like to thank my supervisor Merel Rozier for her endless support, sparring sessions, enthusiasm, and pleasant conversations during this graduation process.

I would also like to thank my supervisors from the university for their guidance and critical feedback. Dorine Duives, as the first supervisor, for the consultations in which she clearly wanted to help me take my thesis to a higher level. Maarten Kroesen for passing on his knowledge about the method I used and the critical view on my work, which I had to get used to at first, but which turned out to be very useful. And Serge Hoogendoorn for his enthusiasm, which always made me feel comfortable during the meetings with my entire committee.

Finally, I would like to thank my family and friends for their endless support during this, rather lonely, graduation process. My roommates, for the support hugs, and interest in what I was working on. And, of course, the joy I could share with all of them when another milestone was reached.

B.D. (Babette) den Hollander
Delft, April 2023

Contents

List of Figures

List of Tables

1 Introduction

1.1. Research context

Crowd formation takes place all over the world. This can be due to an influx of people to a mass event. A mass event is a planned gathering of a large group of people for a common purpose at a specific location (Rodrigues, 2017). Consequently, situations can occur in which too many people come together at the same moment at the same location. This can cause unpleasant experiences or even dangerous situations. To prevent disasters like the crowd crushes in Seoul (Picheta and Bae, 2022) in which 154 people died, from happening, coordination of the movements of people to such an event is essential. When people are spread more evenly over time and space, they will have a more pleasant experience, and money can be saved on security. During COVID-19, the need for the coordinated distribution of people became even clearer. Keeping distance became the norm (Qian and Jiang, 2020).

While designing new buildings or organising events where people gather, safety is one of the most important aspects. Especially at locations where people want to arrive at the same time, it is important to consider the distribution of visitors. People must be directed to prevent crowds during the inflow phase of events. The instructions must be based on the current (or expected) level of crowdedness. However, Kohlhardt et al. (2018) found out that ”visitor satisfaction and crowding perceptions are more likely to be defined by where visitors have these encounters rather than the total number of encounters”. In particular, crowds cannot be indicated with a fixed value. If someone does not experience crowdedness yet, this person will be more difficult to direct. In addition, this crowd perception depends on the visitors' expectations. ”Perceived crowding can be reduced by providing information that makes expectations realistic and allows visitors to select the densities they prefer” (Shelby et al., 1983). By sharing information with travellers, their expectations can be influenced, and as such, their travel choices.

Sweco has developed a tool that aims to predict crowds based on data and tempt people to make different choices. This tool is called Crowd Management System (CROMAS). The consortium of which Sweco is part further consists of Scenwise, Kinisis, TomTom, PTV Group, Vinotion, and Livecrowd. CROMAS consists of three phases: (1) Combining input information (open data, data from clients or third parties, and own data), (2) Predicting crowds on three levels (historical data [a year ahead], short-term data [a few days ahead], and current data [on the day itself, up to a few hours ahead]), and (3) Social media Listening and behavioural influence (informing and advising visitors pre-trip, on-trip and on-site to improve safety, flow, and customer experience). Historical and real-time data are implemented in an interactive dashboard that shows the current and future situation of a specific location or area, making CROMAS the first current, multimodal, and predictive Common Operational Picture (COP). Measures can be taken if this situation does not correspond to the desired situation, such as certain roads becoming too crowded. In CROMAS, these measures consist of triggers, which are informational messages sent via social media, attempting to adjust people's travel choices or messages sent to the organisation to regulate the crowds with, for example, instructions for the traffic controllers (CROMAS consortium, 2022).

1.2. Research gap

To prevent too much crowding at the influx of certain locations, it is important to understand how these situations arise. Studies have focused on mass events to gain insight into how people plan their trips to a location. Furthermore, a lot of research has been conducted into various factors that can influence these trips, known as choice factors. It was found that there are big differences among people. At the same time, little research has been done on the actual applications of these findings during mass events, concerning crowd management. An important reason for this is that mass events often last only one or two days, making it difficult to gather enough reliable and valid data for a study. In addition, to influence the influx of people to a location through information, it is necessary to reach the people who go to the location. In recent years, many studies have used variable message signs as a means of communication, but two drawbacks of this way of communication are that it is not possible to reach all travellers via these signs since these signs are located next to highways, and that travellers can only be reached after departure.

In recent years, technology developed rapidly. Nowadays, almost everyone owns a mobile phone. Most people take their phones with them wherever they go. This has created new possibilities to inform people. These mobile phones can be used to send people travel information. This way of informing people is used in CROMAS. However, the way this information is shared with people is still very general. This is because little is known about the settings of such systems because people's information needs are not fully understood. To improve the effectiveness of CROMAS, CROMAS must be better attuned to the information preferences of different travellers. These preferences consist of the kind of information, the communication channel and the moment they prefer to be approached. In case we know more about these preferences for different types of people, it will be possible to inform people more specifically, especially if the target group of an event is known. Subsequently, it must be investigated whether providing information according to people's preferences will make it easier to influence their travel behaviour.

1.3. Research objective and research questions

This research aims to identify the preferences of different types of visitors towards receiving travel information, and to what extent their preferences match their behavioural intention. Subsequently, these groups can receive more specific travel information that suits their preferences, which aims to influence their travel choices. The long-term goal is to manage the number of people at a specific location at a specific moment by influencing their travel behaviour. Travel behaviour adjustments are defined in this research as changes in modality, route or departure time. This depends on the information contained in the message and what changes are needed at that time. This results in the following research question:

What are the preferences of different types of visitors towards receiving information, and to what extent do visitors with different information profiles act on the information messages they receive?

This question can be divided into several sub-questions:

1. What factors affect modality, route, and departure time choices according to travel behaviour literature?
2. How does travel information influence modality, route, and departure time choices, and what choice factors are involved according to the literature?
3. How can we measure visitors' preferences towards receiving information to update their travel choices?
4. To what extent can different visitor profiles be distinguished with respect to the preferred information provision?
5. To what extent do people with different information preference profiles have different follow-up behaviour?

1.4. Scope

In this study, information refers to travel or event information. The information was provided via messages sent through channels available on mobile phones. Mobile phones present new opportunities for the communication of travel and event information, given how quickly other information can reach people through mobile phones. The selection of communication channels that participants could choose from was based on the researcher's experience and environment, and the options available in CROMAS. The channels include WhatsApp, Facebook, Instagram, LinkedIn, Twitter, Telegram, news apps or websites, event apps (if available), email, and SMS. When people buy a ticket in advance, contact information is sometimes known. For free events, it is mostly unknown who visits the event. Both event types were included in this research.

This research focuses on the influx of mass events in Amsterdam. Mass events are events where more than 500 people gather and can weigh on the environment due to sound pollution, closed routes or modified functions of areas like parking (Heerhugowaard, 2018). Mass events have been chosen as a focus because many people arrive within a few hours, making the influx busy. In addition, research on the influx of mass events is limited, so there is still much to learn. All events included take place in Amsterdam because many mass events in the Netherlands happen in this city. The influx is not always smooth, making this the most interesting location for this research. The three events, music festival, light festival, and football match, were compared. The reason for this selection and the properties of the events are discussed in Chapter 3.

As this study focuses on mobile phones, only studies written since 2009 are included in the literature review because the first WhatsApp messages were sent that year. Social media has become increasingly important in people's pastimes since then. People's preferences regarding receiving the information will vary as different people use social media in different ways. Hence, this study does not focus on any specific group. The participants' preferences that have been studied are related to the content, communication channels, and timing of communication, as these components can be included in CROMAS.

1.5. Relevance

As mentioned in Section 1.2, much research has been done to find out how people plan their trips and based on which factors they make travel choices. This is to limit crowds at locations or the influx to certain locations. Information can be used to influence these travel choices, explained in Chapter 2. But what information, at which moment, and how do you reach the visitors? This research contributes by adding knowledge about the information preference profiles of visitors of mass events in Amsterdam, and the extent to which visitors act on the information messages they prefer. Insight can be gained by using the Latent Class Cluster Model to identify visitor profiles. No other study focusing on travel information preferences using this method was found. However, this information can be important for event organisations and municipalities because they know the target group of the event, and can with these results better align their communication with the preferences of this target group.

The social relevance follows from the user profiles that show that they can be influenced if the information is shared in accordance with their preferences. Less effort is needed for the groups that strongly adhere to their plan. So, better results can be achieved with less effort, and the visitors' experiences will be more pleasant. The results can also be implemented in CROMAS. When the system better meets the information needs of the people it informs, the effect of CROMAS can improve.

1.6. Outline

This chapter provides an introduction to the topic of this research. Chapter 2 (Travel choices) discusses the travel choice factors in more detail, based on literature. A distinction is made between mode, route, and departure time choice factors. The reason why these three categories were chosen can be found in the introduction of Chapter 2. After determining the factors on which travel choices are based, the investigation turned to which factors can be influenced to change travel choices and how this can be

done. This resulted in the conceptual model in which the focus of this research is visualised in figure Figure 2.2 and Figure 2.3. These figures show the relation between aspects involved in this research. Chapter 3 (Research methodologies) explains how the found literature was used to design a survey. Also, the survey dissemination was discussed. Once the data was collected, an Exploratory Factor Analysis (EFA) and a Latent Class Cluster Analysis (LCCA) were performed. A detailed explanation can be found in Section 3.2 and 3.3. This resulted in visitor clusters based on similar behaviour. The results were checked by comparing them with how the participants rated five messages about follow-up probability, to determine whether different clusters would rate these messages differently. The analysis results can be found in Chapter 4, including a description of the participants, general findings, Exploratory Factor Analysis, Latent Class Cluster Analysis, and behavioural intention. Striking general results are highlighted in the first two parts, while in the last two parts, the steps explained in the method are executed, and the clusters are presented. The research ends with a chapter discussing the conclusion, implications, limitations and recommendations for future work.

2

Travel Choices

Travel choices are based on a variety of factors. Complexity arises when it is realised that these factors are not just facts but mostly opinions or just a feeling. Even when the factors are facts, not every person values them equally. This means that these factors can be different for everyone. Furthermore, the factors vary per choice. According to Daamen (2004), the choices that must be made can be divided into three levels: strategic, tactical, and operational. For pedestrians, the strategical level mainly consists of (1) determining the activity choice set, the tactical level consists of (2) activity scheduling, (3) activity area choice, and (4) route choice, and the operational level consists of (5) the choices while moving, such as departing, moving itself, performing an activity and trajectory choice. So, in total, there are five choice types. As this research focused on the influx of visitors to given events, (1) the activity choice set, (2) activity scheduling, and (3) activity area choice were already given, so these choices have already been made. Only the choices (4) route choice and (5) choices while moving (e.g. departure time) still need to be made. Furthermore, multiple modalities were considered by the visitors before departure. Therefore, mode choice has also been taken into account. This resulted in three interesting choices for this research: transport mode, route, and departure time. These three choices each have their choice factors, which are parts or properties that are relevant to decision-making.

This chapter describes the existing literature on choice factors to learn more about how travel choices are made and how they can possibly be influenced. In Section 2.1, choice factors for transport mode, route, and departure time are discussed. After knowing how the choices were made, it was important to see how these choices can be influenced. This is described in Section 2.2. The chapter ends with the conceptual model in Section 2.3, explaining what is still missing in literature and what has been focused on in this research.

2.1. Factors affecting travel choices

The choice factors can be divided into three more general categories based on the results of Hanson and Schwab (1986). Firstly, traveller characteristic factors, which can be defined as personal factors related to the traveller, including socio-demographic factors and behavioural factors. In this study, the term *'personal factors'* is used to describe this. Secondly, the spatial perspective, which can be defined as external factors related to policy, economy, and psychical environment while people are travelling. From now on, this will be referred to as *'external factors'*. And thirdly, factors considering the transport system which relate to the trip itself and can generally be expressed with a number, such as time and cost (Wang, 2015). From now on, this will be referred to as *'trip characteristics'*. With these categories, it appeared that the literature could be described in a way that fits well with this research, where the focus is on the factors that can be influenced. This distribution will be maintained when discussing the choice factors per travel choice below.

2.1.1. Mode choice factors

Much research has been conducted into why people choose a particular mode of transport, with a particular focus on commuters. According to research results of Tyrinopoulos and Antoniou (2012), the availability of parking space was the most critical factor for commuters when selecting their mode of transport, followed by the level of crowdedness. Striking, in this case, was that poor accessibility to the transit network, high fares, and lack of public transport information did not seem to affect the choice. In addition, intention and habit play a significant role in the choice of car travel. "Intentions capture the motivational factors that influence behaviour and are indicators of how hard people are willing to try, or of how much effort they are planning to exert, to perform the behaviour" (Ajzen, 1991). For public transport, only intention plays a role (Donald et al., 2014). The results depend on the location where the research is conducted. The research of Mayo and Taboada (2020) was conducted in the Philippines and showed different results. Regardless of personal factors, safety was ranked as the most important factor. This highlights the local differences in comparison with the Netherlands.

When looking at tourists instead of commuters, other factors become apparent. For example, the number of times they have already visited a place and the costs compared to the price of a taxi. Furthermore, the travel partner and trip motivations have an impact. The motivation can make people prefer to walk to see more of the city (Le-Klähn et al., 2015). In this case, travel distance and infrastructure can make a particular travel mode more attractive. By improving the infrastructure, the government can influence the use of a specific mode of transport (Wang and Liu, 2015). Therefore, it is essential to include leisure travellers in research, as most commuters travel alone, and their trip motivation is already established.

Since the average age of commuters is relatively high, it is also interesting to consider younger people. Young adults have different mode choice factors. This group is often not in possession of a car, which makes this transport mode not an option. Simons et al. (2014) found that autonomy, travel time, financial cost and vehicle ownership were the most important factors for young adults when choosing their mode of transport. Information about health or ecological benefits did not appear to affect their choices. The choice factors of young adults overlap with those of people living in developing countries. This is expected to be related to the limited amount of money available (Toro-González et al., 2020).

An overview of the factors identified is presented in Table 2.1. A + sign indicates that with a higher factor score, someone is more likely to choose that mode. The - sign has the opposite effect. The $\sim$ sign indicates that the effect depends on more factors or the mode of transport. For example, older people are more likely to choose a car as a travel mode but less likely to choose public transport.

The previous paragraphs showed that it is essential to include different types of travellers when analysing mode choice factors. Each group values factors differently when choosing a transport mode. Therefore, Table 2.1 includes several personal factors. It is noteworthy that quite a few external factors have to do with the appearance of the street, which the government can respond to when promoting specific modes of transport. Some external factors, such as weather or perceived safety, are challenging for the government to influence. Another noteworthy factor is autonomy, which is included in the trip characteristics. This factor allows flexibility when travelling. It is a factor that appears less often in the list of trip characteristics, where factors such as travel time, distance and costs are always mentioned.

Table 2.1: Mode choice factors

Personal factors	External factors	Trip characteristics
(∼) Age	(-) Streets with holes, gaps and tram rails	(+) Autonomy
(∼) Financial situation	(-) Narrow paths	(∼) Travel purpose
(+) Vehicle ownership	(-) Poorly maintained paths	(+) Comfort
(+) Habit	(-) Non-existing cycling path	(-) Level of activity
(+) Intention	(-) Unclear road division	(-) Level of crowdedness
(∼) Number of previous visits	(+) One-way streets	(-) Travel time
(∼) Social influence	(+) Accessibility to transit network	(-) Travel distance
(∼) Trip motivation	(+) Availability of parking space	(-) Travel costs
	(+) Access to facilities	
	(-) Weather	
	(+) (Perceived) safety	
	(+) Ecological benefits	

2.1.2. Route choice factors

Basu et al. (2022) attempted to synthesise the existing literature on pedestrian route choice (PRC) and provided an overview of all the factors associated with PRC. This resulted in 105 factors divided into three categories: pedestrian socio-demographic factors, built environment factors, and trip characteristics. These categories correspond to the previously mentioned categories of personal factors, external factors, and trip characteristics. Basu et al. (2022) grouped the 105 factors they found into more general factor groups, which can be seen in Table 2.2 (along with factors found in other literature mentioned in this section). The sign in the table indicates the relation between the factor and the probability that someone chooses a specific route. The $\sim$ sign indicates that it depends on a combination of factors. These factors can influence the route choice positively or negatively (Galama, 2015). It is important to note that not everyone values each factor equally, and therefore, not everyone chooses the same route when exposed to the same set of routes with given route characteristics. These differences between individuals result in distribution over routes.

Table 2.2: Route choice factors

Personal factors	External factors	Trip characteristics
(∼) Age	(∼) Sidewalk characteristics	(-) Distance/length of trip
(∼) Gender	(+) Street crossing facilities	(-) Traffic volume
(∼) Ethnicity	(+) Pedestrian amenities and urban design features along the sidewalk	(-) Walking time and waiting time
(∼) Occupation	(+) Route infrastructure	(∼) Trip purpose
(∼) Income	(∼) Land uses along the route	(-) Vehicle size
(+) Companions	(+) Condition of the buildings	(-) Motor vehicle speed
(∼) Activity pattern	(+) Topography	(-) Number of trucks in the flow
	(∼) Residential and non-residential density	
	(+) Safety	
	(+) Security	
	(+) Quality of walking environment	

To ensure that as many choice factors as possible are considered, it is important to consider different modes of transport. Different factors play a role in different modes of transport. The results of Segadilha and da Penha Sanches's (2014) research into the perception of cyclists concerning route choice factors show that many factors which are important for pedestrians are also important for cyclists. The volume of the traffic, security, and street lighting were found to be important factors. These factors correspond to traffic volume, security, and sidewalk characteristics, which were mentioned by pedestrians. The most important factors pedestrians did not mention were motor vehicle speed and the number of trucks in the flow. It was striking that the personal factors gender, frequency of bicycle use, and age did not show a significant difference between people, while these factors clearly emerged in Basu et al.'s (2022) research.

To obtain a complete overview, the route choices of car travellers were also taken into account and included in Table 2.2. Ramaekers et al. (2013) found that activity patterns influence the road category primarily driven on. In addition, the purpose of the trip showed a significant influence on the deviation from the shortest path. Of course, trip characteristics also play a role. Notably, only a few differences were found between on and off-peak hours, especially for commuters. The utility does not get maximised. In addition to trip characteristics, personal and external factors significantly influence travellers' experiences.

The trip purpose was already mentioned as an important choice factor in the mode choice factors. It is striking that this factor also appears to be important in route choice. Another notable factor that recurs is social influence/occupation. These factors stand out because there is no utility maximisation here, but they still recur in addition to the most common factors in research, such as socio-demographic factors. Furthermore, it is remarkable that in addition to the layout of the streets, safety is also of great importance when making choices.

2.1.3. Departure time choice factors

The choice factors for departure time are divided into the same categories as the transport mode and route choice factors. An overview of these factors can be seen in Table 2.3. The factors mentioned are found in the literature and discussed below the table. A positive sign indicates an earlier departure time and a negative sign indicates a later departure time. The $\sim$ sign indicates that it depends on a combination of factors.

Table 2.3: Departure time choice factors

Personal factors	External factors	Trip characteristics
(+) Early preferred arrival time	(+) Availability of parking spots	(+) Level of crowdedness
($\sim$) Activity schedule	(+) Distance from expressway entrance	(+) Travel time
($\sim$) Companions	(-) Number of tourism spots nearby	(+) Delay penalty
($\sim$) Attitude and norms		($\sim$) Transport mode
(+) Perception toward being on time or cost		(+) Destination
($\sim$) Work field		($\sim$) Trip purpose
(-) Flexibility in work schedule		(-) Flexibility
($\sim$) Previous experiences		(+) Delay
($\sim$) Household		
($\sim$) Age		
($\sim$) Gender		
(+) Experience of being late		

Avoiding congestion is the most important factor for travellers to change their departure time. This was confirmed, among other things, in the research of Thorhauge et al. (2015). Furthermore, people seem to have a specific preferred arrival time. This arrival time can be influenced by a trade-off between travel time and the consequences of being late or early. Other factors that influence the departure time choice are mode of transport, destination, trip purpose, activity schedule, and companions. The level of flexibility is also important, for example, in case of delay or the number of travel options for the return trip. A factor related to preferred arrival time is how people experience arriving late, resulting in less flexibility (Arellana et al., 2012). The importance of these factors can vary greatly from person to person. This is also the case for the experienced level of crowdedness. In addition, more practical aspects are relevant, for example, the number of available parking spots or the travel time (Xue et al., 2019). Furthermore, people take past experiences into account. After experiencing a lot of traffic jams or no free parking spots, people seem to depart earlier to avoid traffic jams.

The trip purpose is again mentioned as a factor that influences choices. This is the case for transport mode, route, and departure time. For commuters, their work field is proving to be of great importance. The degree of flexibility in the work schedule and the result of late entry differs not only by person but

also by work field (Abkowitz, 1981). A lot of research on departure time choices focuses on commuters. However, a significant part of travellers travels for leisure reasons. The departure times for this group are not as flexible as expected due to activity schedules, and individual and household characteristics (Steed and Bhat, 2000). The level of service seemed to be less important. Regarding leisure travel, the travelling group also appears to be of great importance, especially for young children (Le et al., 2022). This travel group also influences other factors, such as the maximum acceptable crowds, travel time, or distance.

In addition to the fact that trip purpose was already mentioned, personal factors such as age and gender, and trip characteristics such as travel time, costs, and distance were mentioned again. It is striking that the level of crowdedness is also often mentioned. Travellers prefer to avoid crowdedness, which shows opportunities for CROMAS. It has also been noted in the previous paragraphs that people like to stick to their own devised departure time. Factors such as how people experience late arrival, and the travelling group reinforced that. External factors appear to occur less frequently as departure time choice factors in the studies analysed.

2.2. Influencing by using informational messages

Now that more is known about the factors involved in people's mode choices (Table 2.1), route choices (Table 2.2), and departure time choices (Table 2.3), the next step is to determine whether these factors can be influenced. A literature review was conducted to investigate how information messages on mobile phones can influence travel behaviour. A flowchart of this literature review process can be seen in Figure 2.1. The search terms used to filter the literature are related to this topic. Table 2.4 shows the search terms, divided into three categories: information, area of influence, and impact. At least one search term from each category had to be found in the title. This resulted in 144 articles. Subsequently, the years in which the studies were conducted were examined. Smartphones have developed at a rapid pace in recent years. The first iPhone was released in 2007 (Verizon Editorial Team, 2023), and the first messages via WhatsApp were sent in 2009 (Pahwa, 2023). Before this time, no information messages could be sent via this platform. Therefore, it was decided to focus on articles published after 2009, resulting in 74 articles.

Table 2.4: Search terms in three categories

Information	Area of influence	Impact
Information	Route choice	Impact
Message	Modality	Effect
Smartphone	Travel mode	Effectiveness
Social Media	Departure time	Reduction
	Travel behavio(u)r	Increase
	Travel choice	

The titles of the articles were read to check if they were relevant to the topic. Specifically, articles in the field of modality effect (as this is a psychological topic) were extracted, leaving 37 articles.

Then, all abstracts were read, and papers which were not accessible or had a different focus area were removed, leaving 20 articles. However, none of these articles focused on influencing departure time. To address this gap, the search for papers on this topic was reconducted, this time without a year limit, resulting in three interesting and accessible articles. These articles, however, did not focus on information messages but rather on pre-trip information or variable message signs. Nonetheless, it was decided to include them to gain insights into how departure time can be influenced.

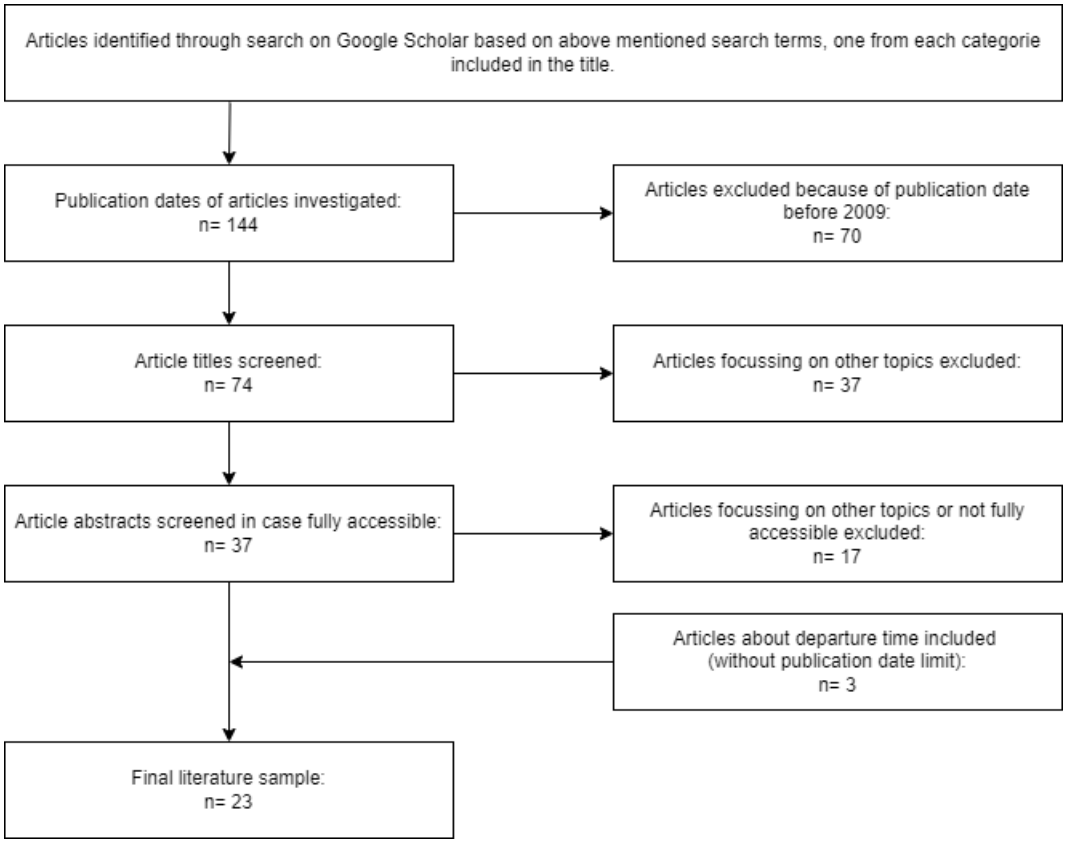

Figure 2.1: Flowchart of literature review process

An overview of the articles can be found in Appendix B. For each paper, the goal, area of influence (such as travel choices or travel efficiency), message type, results, and source are mentioned. Some papers did not specify the travel choices and are therefore described as 'travel choices'. In addition, many studies did not mention what kind of information was displayed and how it was disseminated. In these cases, a more general description is given under 'message type'.

Several things are remarkable in this overview. As can be seen, most studies focus on influencing route choices. This is mostly done by improving the choice factors related to the trip characteristics, limiting the delay and increasing travel efficiency. It is noticeable that most studies focus on car traffic. This is probably because the problems with car traffic are obvious, while the choices made while cycling do not immediately lead to hours of traffic jams. This often makes driving the most interesting mode of transportation to influence, resulting in few studies focusing on limiting travel time or increasing the comfort of people who travel by public transport or by bicycle.

Other notable aspects have to do with the information shared with the travellers. How the information is described is often vague. Some articles refer to 'travel information' or mention the subject or another, not very concrete description of the information given to the participants. At the same time, several studies have shown that the influence of specific content and the form of the information is of great importance. Liu and Guan (2011) and Van Berkum and Van der Mede (1993) found that the influence of the information was related to reliance on the information, in addition to the random level of driver's route choice. Subsequently, Khattak et al. (1995) showed that the perception of messages to be accurate, relevant and timely is of great importance. Furthermore, it takes some time to let people act on the behaviour since people who listen to traffic reports more often are more likely to follow up on the messages. So, the content of the messages must be trustworthy and accurate if you want to change travel behaviour. It is a shortcoming that the information is not clearly described in every study. Further research into communication can therefore be used to improve people's follow-up behaviour.

Only one study used mobile phones to send information messages to travellers to influence their travel behaviour. The results of Gan (2015) show that Smartphone-delivered Multimodal Information (SMMI) can significantly influence mode choice. It was found that the main factors that determine to what extent SMMI can influence mode choice are traveller attributes, driver's previous experience, and level of service attributes (Gan, 2015). These factors were also mentioned in Table 2.2. SMMI does not seem to be used much in research yet, but it appears to be a new and interesting way to influence people's travel behaviour.

It became clear that various factors play a role in follow-up behaviour. Appendix B was converted into Table 2.6 and 2.7, in which lists of choice factors are shown, which appeared to be of significant influence while influencing travel behaviour. The colours indicate the category to which the factors belong. The meaning of each colour can be seen in Table 2.5. The categorisation is in line with the determination of the choice factors in the previous sections, derived from the research of Hanson and Schwab (1986), and explained in the introduction of Section 2.1. The 'message-related factors' category was added because this study focuses on influencing behaviour by sending messages, and the way this is done may also influence the choices. Table 2.6 and 2.7 were used to determine which factors were included in this research. Each category is further explained below.

Table 2.5: Legend for choice factor categories of Table 2.6 and 2.7

Personal factors
External factors
Trip characteristics
Message related factors

Personal factors

Looking at Table 2.6 and 2.7, first, personal factors (in green) were analysed. The most frequently mentioned factors are occupation/education/profession (6 times), age (5 times), gender (4 times), and convenience/habit (4 times). This is consistent with the results of Gan (2015), in which SMMI was used. The study revealed that traveller attributes and drivers' previous experiences influence travel choices. In the choice factors of Table 2.1, 2.2, and 2.3, these factors reappear as financial situation/income/occupation/work field, age, gender, and habit/previous experiences. Gender was only mentioned as a route choice factor, but since not all travel choice factor studies were included and gender was included in the choice factor tables, gender may also influence mode and departure time choice. Another striking factor is social influence or companions, which were mentioned as a choice factor for mode, route, and departure time. This factor was mentioned thrice in Table 2.6 and 2.7 and was also mentioned in Section 2.1.

External factors

The external factors (in red) were analysed second. These factors were not common in the studies. In general, these factors are particularly important when looking at travel comfort and attitudes towards specific modes of transport. Therefore, a significant difference in such a factor is needed to prompt people to change their travel behaviour. In addition, these factors are often not easy to adjust while influencing travel choices. Remarkably, these factors occur much more often in route choices than in other choices. For mode choice, the weather plays an important role. However, people often check the weather themselves, or they are informed by other applications which we cannot change. Therefore, in this study, it was decided not to elaborate on these factors.

An important factor mentioned as a mode choice factor that cannot be easily influenced is access to the network. In addition, (perception of) safety turned out to be an important choice factor, but this is also not easy to influence with information provision and, therefore, cannot be found in Table 2.6 and 2.7.

Trip characteristics

In the trip characteristics category (in yellow), travel/waiting/delay time (8 times) and comfort/crowdedness (5 times) are mentioned most frequently. These factors seem to be important reasons to change plans and, therefore, to influence travel choices. This result also corresponds to the choice factors in Section 2.1.

The analysed studies mainly focused on one specific reason people travel, such as work or leisure. Other studies included various travel motives but looked at different types of people for this (some people travel for work and others for leisure). Therefore, little attention was given to the various trip purposes for the same person. This is striking because trip purpose was mentioned as an important choice factor in Section 2.1. For this reason, the trip purpose was included in this research.

Message related factors

While analysing the message-related factors (in blue), it was found that many studies did not clearly specify how the information was presented to the travellers. Some studies did not even specify which information was included in their messages. As a consequence, the influence of this factor is not always included in the results, which can be seen in the division of the blue cells in the table. Since a few studies have shown that factors related to the content and structure of the message are important, more attention must be paid to the content of the messages during the studies.

Something that did not emerge from the literature, but is important to convey information, has to do with the communication channels. Not every traveller can be reached in the same way, as they each use different media. In addition, people have preferences for certain channels when it comes to travel advice. When people experience messages more positively due to similarity with their preferences, they are less likely to experience the messages as burdensome and are more likely to follow up.

Third, the studies analysed often focus on sending messages when people are already on their way. However, there are more possible moments to send a message, such as before departure or at the location. Perhaps people are much more impressionable at other times. Since this has not much been included in research, this aspect has been included in this research.

Table 2.6: Choice factors travel choices

Area of influence	Message type	Effect	Choice factors
Travel choices	Real-time information	Most popular reaction is change of departure time	- Trip length
			- Age
			- Profession
			- Attributes of information
			- Social norms
	Advanced travel information	Reduction in travel time and potential delays	- Time pressure
			- Landuse
			- Route characteristics
			- Travel time
	Radio traffic information	Follow-up in case of repetition	- Message clearness
			- Explicitly suggested alternatives
			- Age
			- Gender
			- Travel purpose
			- Travel distance

Table 2.7: Choice factors mode, route, and departure time choices

Area of influence	Message type	Effect	Choice factors
Mode choice	Information intervention	Increase use of green travel modes	- Environment awareness
			- Habit of car use
			- Weather
			- Travel time
			- College education
			- Household size
			- Income
	SMMI	Influence mode choice	- Delay for auto
			- Comfort level of rail transit
			- Gender
			- Education level
			- Income
			- Driving experience
			- Driving frequency
			- Main criterion of mode choice
			- Owning an easy public transportation ride card
			- Previous use of P + R
			- Perceived value of existing real-time traveller information
			- Frequency of using real-time traveller information
	Timetable information	Waiting time reduction	- Waiting time
			- Being able to access information any time
			- Location-based services
			- Data integrity checks
			- Cross-device applications
			- Push notifications
			- Personalisation of mobile applications
	PT real-time information	Adaptations	- Waiting time
			- Electric bike ownership
			- Crowdedness in the bus
			- Frequency of using real-time information
			- Age
			- Occupation
	Travel information	Level of effect of choice factors	- Stability
			- Convenience
			- Fare
			- Time
			- Comfort
Route choice	Travel information	Level of effect of choice factors	- Stability
			- Convenience
			- Fare
			- Time
			- Comfort
	Real-time travel information	Travel time reduction	- Potential of real-time information availability
			- Travel time
			- Reliability of the information
			- Network topology
			- Road characteristics
			- Information type
			- Age
			- Driving experience
			- Gender
			- Extend of road congestion
			- Safety
			- Level of regret
			- Information trust
			- Travel activities
			- Number of options
			- Occupation
	Historical data	Enhances behavioural rationality	- Number of intersections
			- Level of crowdedness
			- Habit
			- Route scenery
			- Personal traits
			- Route characteristics
			- Experience
Departure time and route choice	Pre-trip information	More certainty about the outcome of choices	- Occupation
			- Age
			- Gender
			- Number of unaccepted arrivals
			- Absolute schedule delay
			- Message structure
	Real-time information	Majority access, use and respond to information	- Trustful messages
			- Accurate messages
			- Foreknowledge
			- Relevance
			- How many times a message is heard

2.3. Conceptual model

This section discusses how the various components regarding travel choices are related and which components or relations are not sufficiently known yet. This led to the investigation in this research and clarified what needs further investigation. The previous sections showed that travel choices are based on factors that can be divided into four categories: personal factors, external factors, trip characteristics, and message-related factors. Researchers have attempted to influence these travel choices to learn more about controlling crowds. As mentioned in the introduction of Chapter 2, when the focus is on the influx of visitors to a location, the travel choices are about the travel mode, route, and/or departure time.

Some researchers used information messages as a method to influence travel choices. These information messages can be shared with travellers in various ways. This research focuses on communication via mobile phones because this is still a relatively new way of communicating where many people can be reached, since almost everyone owns a mobile phone. This communication method is also implemented in CROMAS. The way the message is presented depends on the message-related factors. Researchers have attempted to influence behaviour by sending messages, resulting in new travel choices. Figure 2.2 visualises how these messages might affect travel choices. The round arrow indicates the reconsideration of the travel choices. The other aspects included in the model are explained below Figure 2.2.

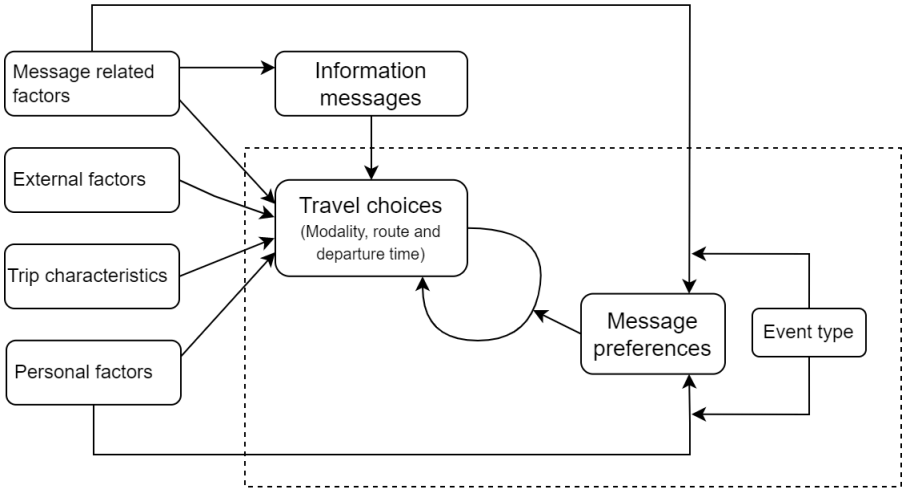

Figure 2.2: Conceptual model, dotted line indicates the focus of this study

It has been shown that not every person processes new information similarly. That is why many choice factors exist, and people's choices are not always easy to predict. This also has to do with people's preferences regarding receiving information messages. The studies analysed in Section 2.2 did not include personal preferences. In those investigations, everyone often received the same general message. However, message personalisation is expected to be important for the extent to which people follow the advice. For this reason, this research focused on the message preferences of the travellers. Personal and message-related factors determine these message preferences since preferences relate to the message presentation options and a person's experiences and character traits.

The previously discussed studies in Section 2.2 demonstrated behavioural differences between commuters and people who travel for leisure (different trip purposes), this has been investigated in many studies for different people, and it was found to be an important travel choice factor. However, one person can differ in their behaviour for each trip purpose. Therefore, this aspect was included in this research. The trip purpose was expected to affect how personal and message-related factors determine message preferences. As mentioned in Section 1.4, at mass events, many people arrive within a few hours, making the influx busy, so it is important that people follow the advice in information messages. Little research has been done on influencing the influx to mass events. Therefore, this

study made a distinction between three different mass events, which is elaborated on in Section 3.1.1. In Figure 2.2, this effect can be seen by the arrows from event type to the arrows from personal factors and message-related factors to message preferences.

People's information preferences need to be known before it is possible to investigate to what extent they influence the reconsideration of travel choices. This research includes a first indication of this relation, indicated by the arrow from message preferences to the round arrow. The dotted line in Figure 2.2 indicates which part of the process is included in this research, including the message preferences of visitors of mass events and the relation between these preferences, and the reconsideration of travel choices.

The area inside the dotted rectangle is shown in more detail in Figure 2.3. The first four personal factors are the most frequently mentioned factors in the studies analysed in Section 2.2. Since these factors showed clear effects in previous research, these factors were also included in this study. How this was done is explained in Section 3.1.1. The last two factors were added to determine whether participants answered questions based on preference or because specific options, such as travelling by their own car, are not for them.

The included message-related factors relate to the presentation of an information message, which depends on its content, communication channel, and communication timing. These aspects are based on previously mentioned studies or have never or rarely been included in previous research. Furthermore, these aspects can be adjusted in CROMAS. This study investigated the effect of these message-related and personal factors on the message preferences of different people and whether the event type influences these preferences.

Moreover, it is noticeable in Figure 2.3 that there is a dotted line from the message preferences to the reconsideration of travel choices. This line represents the second part of the research question, the extent to which visitors with different information profiles act on the information messages they receive. The extent to which this relation exists is not clear yet, and that is why this relation was investigated in this study. This resulted in a first indication of the link between message preferences and reconsidering travel choices.

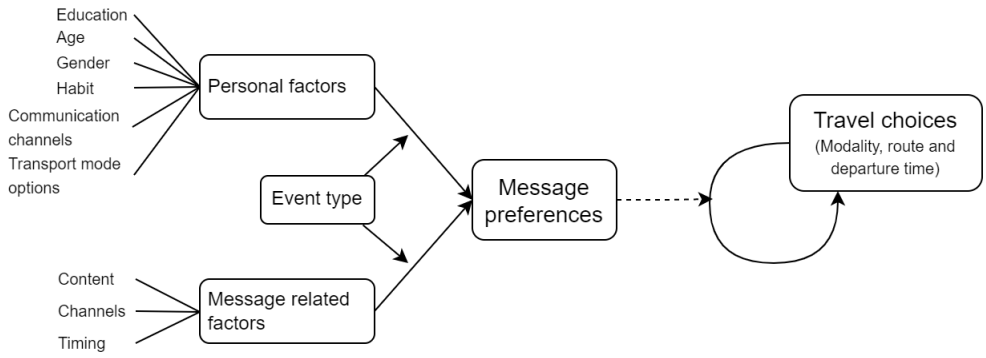

Figure 2.3: Zoomed conceptual model, focus of this research

3

Research methodologies

Chapter 2 showed that little is known about influencing travel behaviour by sending information messages people receive on their mobile phones. To gain more insight into how to influence travel behaviour by sending information messages via mobile phones, a quantitative study was conducted, for which a survey was distributed. This chapter discusses why a survey was chosen, how this survey was set up, and how the data were analysed.

To begin, the aspects of the survey were discussed, including its structure and dissemination in Section 3.1. Next, the way in which the data were analysed is elaborated step by step, including an Exploratory Factor Analysis in Section 3.2 and a Latent Class Cluster Analysis in Section 3.3. Finally, the way in which the results were evaluated by comparing them with the message ranking scores is explained in Section 3.4.

3.1. Survey

There are two types of research, qualitative and quantitative. Qualitative research is used to understand concepts, thoughts or experiences, whereas quantitative research is used to confirm or reject theories and hypotheses. In this study, possible relations were investigated, which made a quantitative approach more appropriate. Common methods are surveys and experiments. A large response group was needed to perform the analysis. A survey worked well for this research because it allowed a large response group to be generated. Surveys typically result in more objective and generalisable results than interviews (Harris and Brown, 2010), increasing the chance that the results give a good first impression of the effects. If a physical experiment was chosen, the probability that clear connections would emerge within the time frame of this study would have been smaller, as the experiment would have needed to be performed several times to rule out coincidence. Besides that, it would be difficult to measure information message preferences in a physical experiment.

The survey was designed using Qualtrics. This platform is more secure than many other options, and privacy is guaranteed to a larger extent. It is an user-friendly platform, available via TU Delft and offers many options for setting up questions. This tool also allows the survey to be completed on a mobile phone, which is very important for this research because of how the survey was distributed. The format, therefore, compares well with other formats according to various comparison sites.

It was decided to distribute the survey in Dutch because most visitors of events in Amsterdam are expected to have Dutch as their native language and live in the Netherlands. By allowing the questionnaire to be completed in their native language, the questions are expected to be well understood, and more people are willing to complete it. Having a survey in two languages increases the risk of differences in the translation and interpretation of the questions, which was not desirable for the expected small number of people who would fill it out in English.

3.1.1. Survey structure

The survey design is based on the conceptual model presented in Figure 2.2. As illustrated in this model, this research assumes that messages-related and personal factors influence message preferences, which are influenced by the event type. Subsequently, these preferences influence travel choices. These assumptions informed the survey structure, which consists of three parts: event-related questions, behavioural intention, and personal factors. This structure is visualised in Figure 3.1, and a detailed explanation of each part is provided in the following sections. In the first part, a distinction is made between the events where the preferences have been asked. The participants are divided into clusters based on these answers. The overlap between the indicated message preferences and the follow-up behaviour was tested in the second part. The third and final part includes questions about personal factors. This is the final part of the survey, as research has shown that people are more likely to complete the survey if the final questions are not too challenging. By this point, they have already spent time and effort answering the questions, which makes them eager to complete the survey (TopScriptie, 2022).

The factors that have been included are derived from the knowledge obtained in Chapter 2 and are further discussed in the sections below. The survey comprises 22 questions, including multiple-choice, 5-point Likert scale, and matrix table questions. The survey does not include any open-ended questions (except for age). This ensures that the survey is easily accessible, straightforward to complete, and easier to analyse because the number of different answers is limited. A disadvantage of this is that the participants were forced to select only the options that were given. This can lead to bias in the results. The complete survey can be found in Appendix C.

The survey started with an introduction explaining the purpose of the research and the investigated subject. Additionally, the survey duration was indicated, and it was emphasised that participation is anonymous and that the information provided will be treated with confidentiality. The researcher was also introduced, and the participants were thanked in advance for their participation.

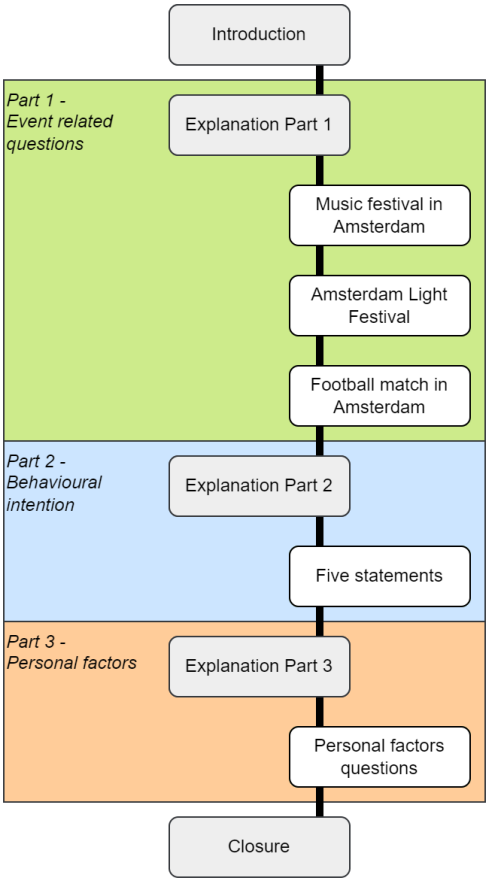

Figure 3.1: Survey structure

1. Event-related questions

The first goal of this research was to learn more about the preferences of different types of visitors towards receiving information. Figure 2.2 showed that the event type is expected to influence the preferences determined by message-related and personal factors. This was included in this study by asking some questions multiple times, each time for another event type. The questions focused on preferred travel mode and information provision preferences. Therefore, all arrows that lead to 'Message preferences' in the conceptual model have been examined here.

A mass event is often described by some practical aspects, including location, date, price, and time (Festival Fans, 2023). Events that differ based on these three aspects have been chosen to see the influence of the various aspects on information preferences. Attention has been given to the fact that the events differ in the target group to be able to include the habit aspect for everyone. The events are all located in Amsterdam to keep the later discussed aspect of familiarity in Amsterdam (which is related to habit) the same for all events. A mass event with a fixed start time and free entrance could not be found and is therefore not included in this research. To limit the number of survey questions, it was decided to investigate three different events, which differ in how often it occurs, price, and start time. An overview of the characteristics of the chosen events can be found in Table 3.1. To ensure that

all participants had the same definition of the three events, a short description was provided before the questions regarding a particular event. The descriptions included the practical information of Table 3.1. The description of the music festival is given below. The other descriptions can be found in Appendix C.

Stel je bij het beantwoorden van de volgende vragen voor dat je het volgende evenement bezoekt:

***Muziekfestival in Amsterdam**, een grootschalig, betaald evenement (meer dan 5.000 bezoekers) waarvoor je van tevoren een kaartje hebt gekocht. De tijd waarop je gaat mag je zelf bepalen.*

Table 3.1: Event characteristics

Event	Occurring	Ticket	Start time
Music festival in Amsterdam	Not often	Ticket	Free entry time
Amsterdam Light Festival	Once a year	Free	Free entry time
Football match in Amsterdam	Often	Ticket	Fixed time

The questions asked for each event are based on the conceptual model in Figure 2.3. The message-related factors were included in the survey by asking about the message content (Q4, Q9, Q14), trip preparation channels (Q3, Q8, Q13), and preferred time to receive information (Q5, Q10, Q15).

The preferred communication channels were only asked once (Q16), as they depend on the social media (SM) channels used by the participants, which is personal and not expected to change per event. A list of options is provided for each question, which can be found in Appendix C. If multiple answers were possible, this was indicated. The previously mentioned personal factor, habit, can also differ per trip purpose, including travel mode habit or preparation timing habit. Therefore, questions about the preferred mode of transport (Q1, Q6, Q11) and trip preparation timing (Q2, Q7, Q12) were included.

2. Behavioural intention

The second part of the research aimed to investigate whether people who expressed a preference for receiving certain types of information would actually act on that information, which was the second goal of this research. It was examined whether the different groups of visitors, each with their own message preferences, also respond differently to the information. This is visualised in Figure 2.2 by the arrow from 'Messages preferences' to the reconsideration of the travel choices. Five messages were presented to the participants based on departure time, route, mode of transport, parking availability, and public transport timetable. These subjects were chosen to investigate whether follow-up behaviour differs depending on the subject matter. The messages are messages which were, if necessary, ready to be sent during a previous CROMAS experiment. One message was chosen for each of the five topics. The participants were asked to rate the messages on a 5-point Likert scale based on the likelihood that they would change their plans based on the information provided. The scale ranged from 1 (absolutely not) to 5 (definitely). The messages are:

1. *Het is momenteel erg druk in de bus van 14.00 uur richting het muziekfestival in Amsterdam. Reis mits mogelijk mee met een volgende event bus. Bekijk hier de locatie van de bushalte en de actuele aankomst- en vertrektijden.*
2. *In verband met een verkeersongeluk vormt zich een file op de toegangswegen naar het muziekfestival in Amsterdam. Klik hier voor de nieuwe snelste route en voorkom files.*
3. *Kom niet meer met de auto naar het muziekfestival in Amsterdam! Er zijn veel vertragingen op de wegen - en de parkeerplaatsen zijn overvol. Ook elders in de stad is geen parkeergelegenheid meer beschikbaar. Wil je toch nog naar het muziekfestival in Amsterdam? Kies dan voor het Openbaar Vervoer of de fiets.*
4. *De parkeersituatie is aangepast als het muziekfestival in Amsterdam plaatsvindt. Bespaar tijd en boek je je parkeerticket vast online!*
5. *De event bus van station Amsterdam Centraal naar het muziekfestival in Amsterdam van 15.00 uur is uitgevallen. Wij voorspellen dat er meer reizigers bij de bus stop bij Amsterdam Centraal zullen zijn dan dat er in de bus passen. Neem een bus later! De volgende bus vertrekt om 15.30 uur.*

3. Personal factors

The questions in this part of the survey are based on the personal factors mentioned in Figure 2.3. In Section 2.2, several personal factors were found to significantly influence travel behaviour choices in previous studies. These factors included occupation/education/profession, age, gender, convenience/habit, and companions. It was expected that these personal factors also influence message preferences. In the conceptual model, this is indicated by the arrow from 'personal factors' to 'message preferences'. These five factors were included in the survey. For education level (Q20), the participants were asked to name their highest completed level of education, which they could choose from six different categories. Age was requested as an integer (Q18), which makes it possible to determine the categories during the analysis. For gender (Q19), participants could choose from four categories. Besides male and female, the options were 'other' and 'I don't want to say'. Since it is not possible to ask directly if behavioural choices are based on habit, the participants were asked if they were familiar with Amsterdam (Q21). When people travel to a place where they are familiar, their interest in travel advice might be influenced by habit. The factor companions was included in the message-related questions about the trip preparation timing. One of the answer options is that the companions prepare this trip. It has not been asked separately whether people travel alone or in company because this research focuses on the preferences regarding receiving information. By asking if people travel alone, nothing is revealed about their preferences.

In addition, a question about the participants' mode of transport options was included (Q22). This was to determine if they make specific choices because they have no other options or if the mode of transport is their preference.

3.1.2. Survey dissemination

Before the survey was distributed, the questionnaire was checked for ambiguities by having it completed by people within and outside the working field. The feedback was processed before the survey was publicly distributed. Furthermore, ethical considerations were considered, and participation was voluntary and anonymous, as stated in the introduction. The combination of personal factor questions cannot lead to the identification of one specific person.

For the survey, a minimal sample size of 300 was desired to ensure the model could fit and function adequately with a Latent Class Cluster Analysis (Nylund-Gibson and Choi, 2018). The suitability of the sample was then also checked using a number of tests, explained in Section 3.2. The survey was shared online among a convenience sample and their networks, through SM platforms such as LinkedIn, WhatsApp, and Facebook. The survey was distributed among employees of Sweco through the department's mailing list and shared LinkedIn posts.

Online distribution enables the collection of many responses in a short time, and sharing the survey with personal and professional networks increases the likelihood of participation since people are more willing to complete a survey of someone they know than someone they don't know. However, this method also has drawbacks, as the backgrounds of the participants reached in this way may not be diverse. Many participants are highly educated and live in specific regions of the country. There is also little diversity in cultural background, partly due to the survey being available only in Dutch. The detailed sample characteristics were described in Section 4.1.

The survey was available from December 15, 2022, to January 18, 2023, and most people were contacted in the first week. Several reminder messages were sent during the rest of the period, and people were verbally asked to complete the survey. Additionally, during this period, it was managed to distribute the survey link in Facebook groups that connect people interested in events in Amsterdam.

After the survey was closed, responses that were not fully completed were excluded from the results because the personal factors were unknown, ensuring the personal description could not be completed. Results where other answers were missing did not exist because all questions were mandatory to answer. Based on the answers to the answer option 'other', it was checked whether there were people who had not seriously completed the survey and had to be deleted, this was not the case. The results were then converted into numbers instead of answers. The results were then converted into numbers, with answer A becoming 1 and answer B becoming 2. After the data was cleaned up, the analysis began.

3.2. Exploratory Factor Analysis

A factor analysis was conducted to reduce the number of variables by combining related ones, and to gain insight into the underlying theoretical structure of the data (Statistics Solutions, 2021). The factor analysis is based on the idea that there are deeper underlying factors that connect the variables. There is a trade-off between detail and the ease of working with data. The best factor analysis solution is the one that combines the variables, minimising the loss of accuracy (Qualtrics, 2022).

There are two types of factor analysis, Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA). EFA is used to identify the underlying factor structure of a set of observed variables without proposing a structure for the outcome. CFA is used to validate a suggested structure of a set of observed variables. With CFA, the hypothesis about the relationship between variables and their underlying latent construct is tested (Suhr, 2006). The different factor analyses are illustrated in Figure 3.2. Here, the f's represent the different factors determined by the factor analysis. The y's are the observed variables; in this research, the answers to survey question(s), for example, different message content. The e's represent the measurement errors, which indicate to what extent a participant answers unpredictably, resulting in fluctuations in the data collection. With EFA, the observed variable can potentially be a measure of every factor, whereas, with CFA, the observed variable can potentially be a measure of only one factor (Columbia University Mailman School Public Health, 2014), indicated by the arrows from the f's to the y's. Another significant difference is that CFA is often used for factors that have been tested before, whereas EFA is often used for factors that have not been tested before (Osborne, 2014).

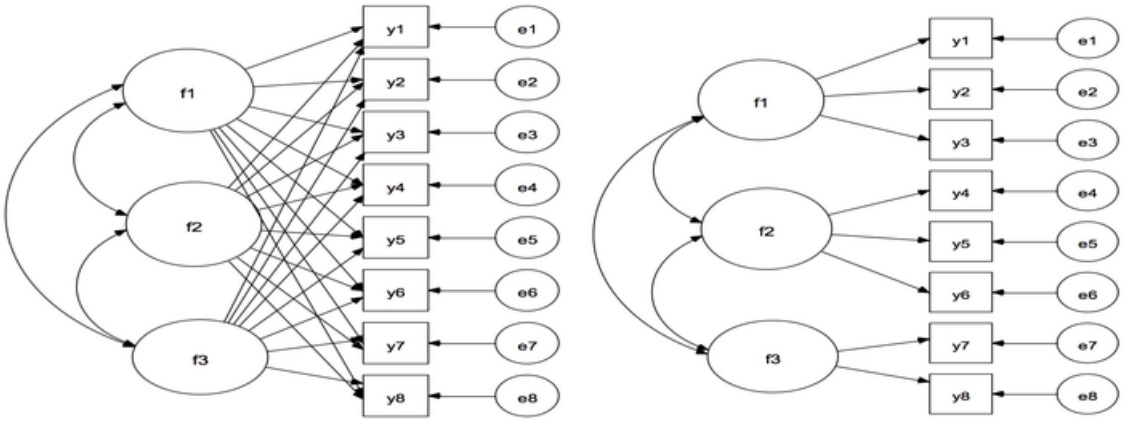

Figure 3.2: EFA (left) and CFA (right) (Columbia University Mailman School Public Health, 2014)

Since no previous research was found that examines the information message preferences of visitors of mass events, there is no hypothesis to test. In addition, the observed variable (survey question answer) can potentially be a measure of every factor (visualised as f in Figure 3.2), making EFA the most suitable method for this study.

EFA assumes that observed variables are influenced by common factors (which influence multiple observed variables), specific factors (which influence only one observed variable), and the measurement error (due to unsystematic events that influence measurement) (Columbia University Mailman School Public Health, 2014). Both common factors and specific factors are latent variables. EFA should not be confused with the principal components analysis (PCA) reduction technique. The main difference with the principal components analysis (PCA) is that with PCA, the goal is to explain as much of the total variance in the observed variables as possible, while with EFA, the goal is to explain the covariance between variables. The observed variables are defined as linear combinations of the factors (Neill, 2013).

The performance of an EFA consists of five steps, which are visualised in Figure 3.3. Each step is explained in more detail below. The EFA was performed using IBM SPSS Statistics (version 28).

3.2.1. Step 1: Is the data suitable for factor analysis?

To begin with, the sample size was assessed. Researchers have different opinions on the minimum sample size needed to perform a factor analysis with a statistically significant result. For example, Tabachnick et al. (2013) suggested a minimum sample size of 300, while Hair et al. (1995) stated that the sample size should be greater than 100. Schreiber (2017) proposes another definition, stating that the minimum sample size is related to communality, and a smaller sample size is needed if the communalities are higher. These disagreements arise from the complex dynamics of the factor analysis. This study initially assumed a minimum sample size of 300, as this value emerges in many studies (or a lower value, but 300 is also sufficient in that case).

Subsequently, the obtained sample size was evaluated based on the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy and Barlett's Test of Sphericity (Williams et al., 2010). These values provide information about the suitability of the respondent data for factor analysis. The KMO measures sample sufficiency for each variable in the model and for the complete model. The KMO value ranges from 0 to 1, with 0 indicating a dispersed correlation matrix and 1 indicating a dense pattern of correlations (Schreiber, 2017). Barlett's Test of Sphericity compares the correlations between the observed correlation matrix and the identity matrix. Tabachnick et al. (2013) and Hair et al. (1995) did not agree on the minimum sample size, but they did agree on the minimum value for KMO and the p-value for Bartlett's Test of Sphericity. They both consider 0,5 as a suitable KMO value and a significant value for Bartlett's Test of Sphericity ($p<0,05$), indicating at least one significant correlation between two variables.

The final values to check to ensure that the data is suitable for factor analysis are the extractions of the communalities. It is desirable that the values are higher than 0,3 (Field, 2009, Allant International University, 2015).

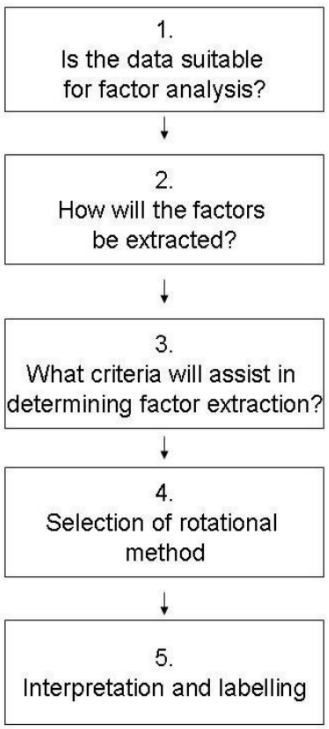

Figure 3.3: Steps to perform an Exploratory Factor Analysis (Williams et al., 2010)

3.2.2. Step 2: How will the factors be extracted?

The second step is to choose a factor extraction method to obtain the eigenvalues, which indicate the total amount of variance that can be explained by an underlying variable or construct. The factor extraction method is used to determine how many factors should be extracted from the observed data (Field, 2009). There are several ways to extract the factors, with Principal Components Analysis (PCA) and Principal Axis Factoring (PAF) being the most commonly used methods in the literature (Henson and Roberts, 2006). PCA focuses on maximising the inter-individual variance (Schreiber, 2021), meaning that it tries to create an index where people differ the most. In PAF, the original correlation matrix and factor loadings are used to determine the communalities. This technique replaces the old communalities with new ones until the values are very small or a set criterion is met (Schreiber, 2021). Thompson (2004) observed that many statistical programs default to PCA, and therefore it is the most commonly used method in EFA. Furthermore, PCA is recommended when there is no prior knowledge (Gorsuch, 1983). This method was also used in this research.

3.2.3. Step 3: What criteria will assist in determining factor extraction?

The aim of the factor analysis is to limit the number of variables by combining them into factors. To determine the number of factors, several criteria are available. Thompson (2004) stated that "simultaneous use of multiple decision rules is appropriate and often desirable", so multiple criteria are needed. Most researchers perform multiple measures (Hair et al., 1995). Popular criteria, which many researchers agree on, are Kaiser's criteria, the scree test, and the parallel analysis (Williams et al., 2010). These criteria were also tested in this research to determine the number of factors.

The Kaiser rule evaluates the eigenvalue of the total variance explained matrix, and the eigenvalues of factors need to be higher than 1 to be retained (Field, 2009). According to the Kaiser rule, the number of factors with an eigenvalue above 1 is the number of factors over which the variables will be distributed.

However, the Kaiser rule is not always reliable, so the results need to be checked (Schreiber, 2021). The scree plot is used to check the results, where the eigenvalue and the component number are plotted against each other to visualise the amount of variance explained by each factor. The number of factors is determined by looking at the component number just before the 'bend' is reached (Columbia University Mailman School Public Health, 2014).

However, this test sometimes shows an unclear number of factors, which is why parallel analysis is also commonly performed. Thompson (2004) even stated that "parallel analysis appears to be among the best methods for deciding how many factors to extract or retain". This analysis compares the eigenvalues with random order eigenvalues, and the actual values need to be bigger than the 95 percentile values of the parallel analysis table. The values for the parallel analysis were obtained from: *https://www.statstodo.com/ParallelAnalysis.php*.

3.2.4. Step 4: Selection of rotational method

Factors are geometrically rotated to make it easier to interpret the results. Rotation maximises high item loads and minimises low item loads, resulting in a more interpretable and simplified solution (Williams et al., 2010). There are two types of rotation: orthogonal, where factors are not allowed to be correlated, and oblique, where factors may occupy any position in factor space and can be correlated with each other (indicated by component correlations close to 0.5) (Columbia University Mailman School Public Health, 2014). For orthogonal rotation, the Varimax method is often used, and for oblique rotation, the direct Oblimin method is mainly used (Kootstra, 2004). If the orthogonal method does not show enough correlation, it can be assumed that the oblique method, in which the factors do not correlate, is appropriate for the data (Allant International University, 2015). The settings for both methods can be found in SPSS.

3.2.5. Step 5: Interpretation and labelling

During interpretation, it was examined which variables were grouped as a factor. It is important that the factors can be substantiated theoretically. A factor should include the loading of at least two or three variables to result in a meaningful interpretation (Henson and Roberts, 2006).

Before investigating the theory behind the factors, the reliability of the results was examined. For this, the Cronbach's Alpha values of all factors were examined. Higher values indicate more overlap between the variables included in that factor (Straub et al., 2004). So, person X consistently answers the questions aggregated as a factor in the same way. According to Straub et al. (2004), the minimum value for Cronbach's Alpha is 0,6. It is often stated that values below 0.6 are too low, values between 0,6 and 0,8 are moderate but accepted, and values between 0,8 and 1,00 are very good (Daud et al., 2018).

3.3. Latent Class Cluster Analysis

The Latent Class Cluster Analysis (LCCA) was performed to find groups with similar message preferences. The LCCA was executed in LatentGOLD v5.0 after the number of variables was reduced by performing EFA. The LCCA statistically identifies latent classes within a population where people in the same class have similar preferences. In other words, it aims to identify 'hidden groups' within the population. According to Collins and Lanza (2009) a latent class is defined as:

- "**Latent** implies that the analysis is based on an error-free latent variable"
- "**Classes** are groups formed by uncovering hidden (latent) patterns in data"

The LCCA identifies these latent (unobserved) classes in the data. "The goal is to maximise homogeneity within clusters and heterogeneity between clusters" (Kroesen, 2022). It is important to note that this method assigns individuals to classes probabilistically, indicating the probability that an individual belongs to a particular class (Molin et al., 2016). Therefore, it is not a fact but a probability. This method can be used when a variable is not easy to measure, such as an attitude, or when there is a topic where people often do not answer questions honestly.

There are multiple advantages of Latent Class Models compared with traditional models. Firstly, no type of distribution needs to be assumed, whether linear or normal, etc. Secondly, during the analysis of the different classes, the external variables (covariates) can be included simultaneously. Thirdly, in the Latent Class Analysis, mixed shell types can be used in the same analysis, including nominal, ordinal, continuous, and/or count variables (Vermunt and Magidson, 2005).

To start with, a model with only indicators was estimated. Indicators are used to assign people to classes. This research focused on the message preferences of different visitor groups. The message preferences consist of message-related factors since these define the preferences. As can be seen in Figure 2.3, the message-related factors consist of the message content, communication channels, and communication timing. Since these aspects define the message preferences, they are the indicators in the model. These indicators are based on the latent variable, which in this study is the attitude towards information messages. This is visualised in Figure 3.4 as the measurement model. The latent variable takes the association between the indicators into account, and it is assumed that the indicators are independent of each other. Some of the indicators are the factor scores which resulted from the EFA.

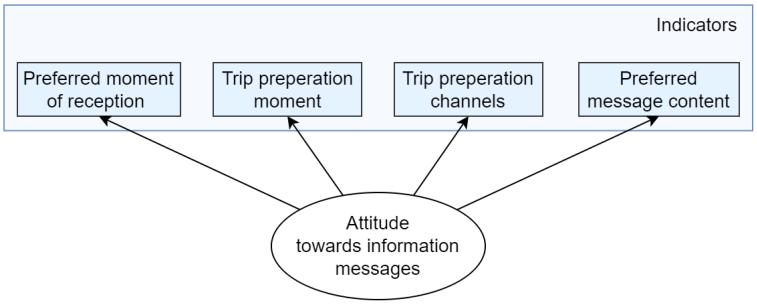

Figure 3.4: Representation of measurement model, the first step of the Latent Class Cluster Analysis *(based on (Volberda, 2020))*

The measurement model is used to determine the optimal number of classes. To do so, three criteria were checked, namely the Bayesian Information Criteria (BIC), Akaike Information Criterion (AIC) and Bivariate Residuals (BVR) (Molin et al., 2016). The value for both BIC and AIC needs to be as low as possible since this indicates the most parsimonious model (Vermunt and Magidson, 2005). When the values get smaller with each increase in clusters, the percentage change in the BIC value from the previous cluster model can be used to determine the optimal number of classes (van't Veer et al., 2023). The optimal number of clusters has been reached when the percentage change represents only a small improvement in model fit. For the BVRs, a maximum value of 3,84 is allowed. This is because values above 3,84 are statistically significant at a 5% level (Schreiber, 2017). This is undesirable as it would imply a significant relationship between the indicators. In that case, the local independence assumption is not held, which is a requirement of the LCCA (Uebersax, 2000). If the requirements are not met, the number of indicators will have to be reduced until a suitable model is found. This is a trade-off between model complexity and fit, and the model that strikes a balance between the two was chosen.

After the number of clusters was determined, the covariates were added to the model, which describe the characteristics of the individuals in that class, also called observed variables, such as age and income. In this study, the covariates are the personal factors that influence the message preferences in the conceptual model in Figure 2.3. These factors do not define the message preferences but might affect them, which is exactly what a covariate does. The habitual question regarding the preferred mode of transport was asked for all three events and therefore returned three times as a covariate. This resulted in the structural model, visualised in Figure 3.5.

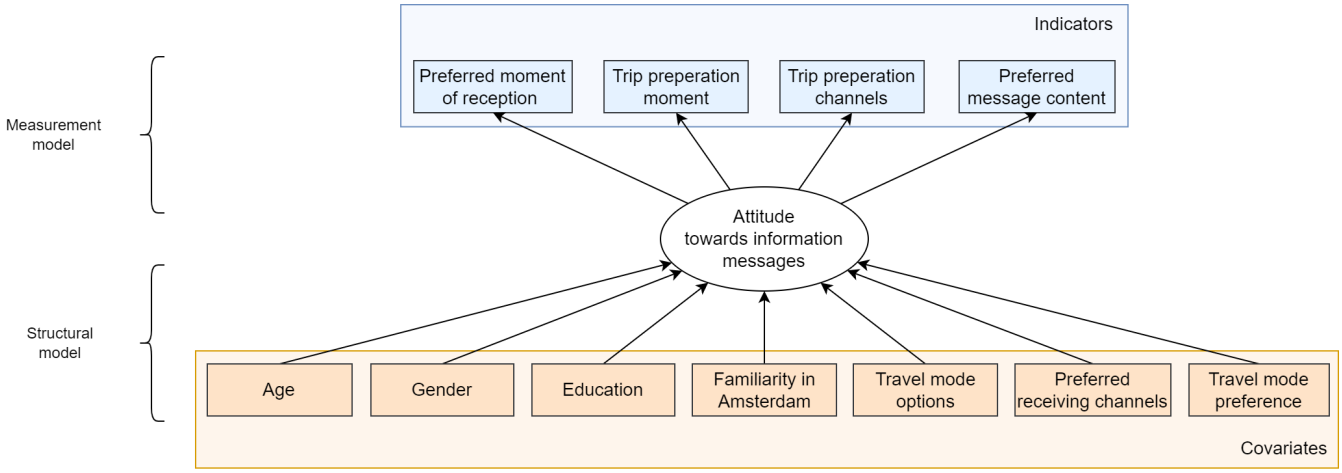

Figure 3.5: Representation of the measurement and structural model second step of the Latent Class Cluster Analysis *(based on (Volberda, 2020))*

The Wald value is used to determine the model fit. Wald values below 3,84 or p-values above 0,05 indicate insignificant covariates, which are not desirable because only a significant relation means that the covariate is likely to predict the class membership in the population (Kroesen, 2022, Knijn, 2020). Backward elimination was used to determine the optimal combination of covariates. The elimination was applied until no more Wald values were lower than 3,84 or p-values higher than 0,05 were found. Backward elimination consists of multiple steps (Chowdhury and Turin, 2020). First, all covariates were added. Not all covariates were significant, so action had to be taken. In the second step, the model was estimated again, but this time leaving one of the covariates out. This was done for all covariates. Third, the covariate with the smallest effect on the model fit was eliminated. This was measured by the decrease in the BIC value. The highest positive value or, when there were no positive values, the smallest negative value was eliminated by making it inactive (van't Veer et al., 2023). Inactive covariates do not affect class membership but can be used to describe the classes. In step four, the Wald and p-value were reviewed again. If not all values were significant, steps two and three were performed again. These steps were repeated until all remaining covariates were significant.

After the significant combination of covariates had been found, it was checked how accurately the model calculates the clusters in combination with the covariates (Vermunt and Magidson, 2005). This is indicated by the entropy R-squared. This is a value between 0 and 1, where 0 indicates a very bad prediction, and 1 indicates a perfect prediction (DiStefano and Kamphaus, 2006). Clark and Muthén (2009) mentioned in their research for an entropy of 0,8: "It was most likely class membership was the best performing method in terms of recovering the true value used in the simulation study, and had relatively good coverage and power in the settings examined".

3.4. Behavioural intention

After determining the preferences of different groups of visitors towards receiving information, it was investigated whether these groups reacted to the information in a different way, indicated by the dotted line in the conceptual model in Figure 2.3. Does a preference for a specific type of information actually translate into action on the information message? It is important to keep in mind that all participants were probabilistically assigned to a cluster, which means that they do not always have to belong to the same cluster. According to Bayes' rule (Bakk et al., 2014), the posterior probability of belonging to a cluster t is given by:

$$P(X_i = t|Y_i) = \frac{P(X_i = t)P(Y_i|X = t)}{P(Y_i)} \tag{3.1}$$

In words, this means that the conditional probability of an answer coming from cluster t, which is part of population X_i, given that the answer is Y_i. The answer Y_i is assigned to the cluster from which the value of P is the greatest. It is most likely that this answer is given by this cluster.

The model assignment rule was used to generate posterior classification, which is the most widely used method (Collins and Lanza, 2009). Here, participants are assigned to a cluster based on the highest probability, leading to 'hard partitioning' (Bakk et al., 2014).

After the participants were assigned to a cluster, the ratings given to the messages of the second part of the survey were analysed per cluster. The messages can be found in Section 3.1.1. The distributions of the different scores among the clusters were compared. It was examined whether there were remarkable differences between clusters and whether clusters gave a higher ranking for information in which they indicated an interest. This gave a first indication of the actual behavioural intentions. Since only five messages were shown to the participants, no conclusions could be drawn from this, but striking observations may lead to future research.

4

Results

This chapter describes the results obtained from elaborating on the method presented in Chapter 3. The data was obtained through the survey described in Section 3.1. The results are divided into five parts. Firstly, the personal factors of the participants are described in Section 4.1. Secondly, the data set's general and noteworthy results are elaborated on in Section 4.2. After that, the steps to perform EFA and LCCA, explained in Section 3.2 and 3.3, were executed. The results from this are described in Section 4.3 and 4.4. Finally, the results of the behavioural intention per cluster are compared in Section 4.5.

In total, 539 people clicked on the link to the survey, and 378 of them completed the survey, resulting in a response rate of 70%. Incomplete responses were filtered out since they cannot be used to describe the clusters (the personal factors were included in the last part of the survey). The sample size of 378 is more than the previously mentioned minimal sample size of 300 (Section 3.1), which meant that the EFA and LCCA could be performed. The average time taken by the participants who finished the survey to complete it was 36.5 minutes, which is long, possibly because some responses were open for several days. The median time taken was 7.1 minutes, which gives a better idea of the time it took to complete the survey.

4.1. Descriptive personal results

To determine the extent to which the sample is representative of the Dutch population, the personal factors were compared with data from Statistics Netherlands (CBS, 2022, 2023). The percentile distribution of participants was compared with that of the Dutch population for age, gender, and the highest level of education completed. These comparisons are visualised in Figure 4.1, where the first column shows the study participants and the second column shows the distribution within the Netherlands.

Table 4.1 presents an overview of this data, along with the other queried personal factors, familiarity with Amsterdam and travel mode options. It should be noted that there are no available values for these factors within the Netherlands. The significance of the differences between the participants and the Dutch population was calculated based on a confidence level of 95%. An asterisk in Table 4.1 indicates a significant difference at this confidence level compared with the Dutch population.

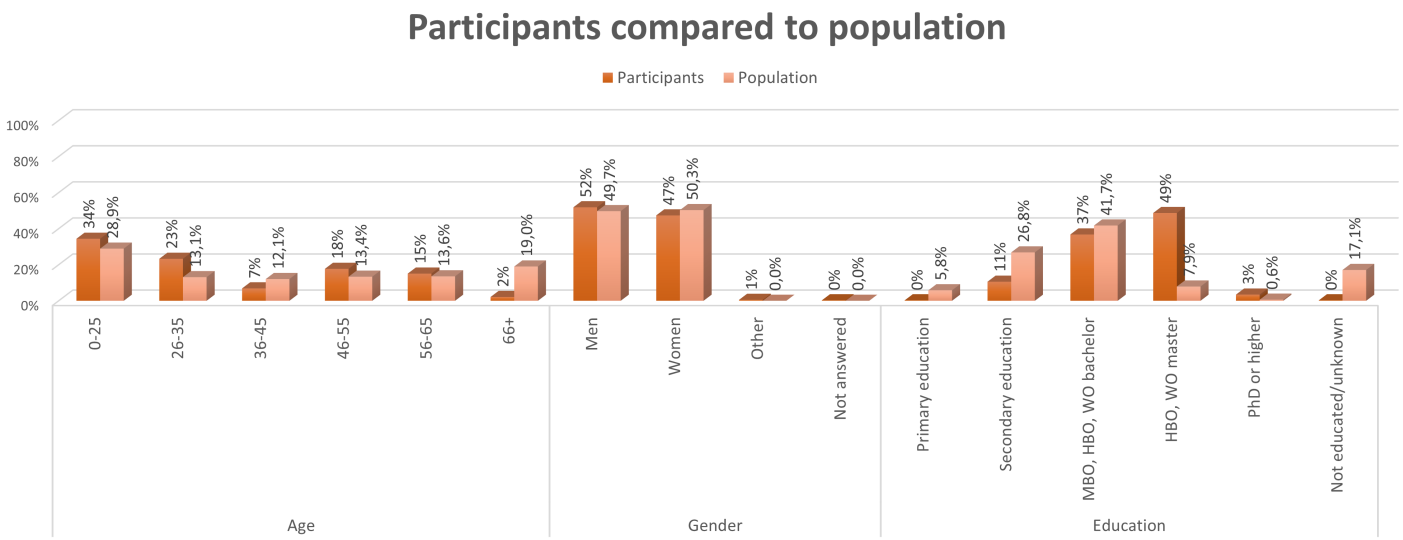

Figure 4.1: Personal factors age, gender, and education level of the participants compared with Dutch population (CBS, 2023, 2022)

Table 4.1: Distribution of personal factors within the participants and the Dutch population (CBS, 2023, 2022) (* means a significant difference at 95% confidence level)

		Participants	Population
Age	0-25	34,4%*	28,9%
	26-35	23,3%*	13,1%
	36-45	6,9%*	12,1%
	46-55	17,7%*	13,4%
	56-65	15,1%	13,6%
	66+	2,1%*	19,0%
Gender	Men	51,9%	49,7%
	Women	47,4%	50,3%
	Other	0,5%*	0,0%
	Not answered	0,3%	0,0%
Education	Primary education	0,3%*	5,8%
	Secondary education	10,6%*	26,8%
	MBO, HBO, WO bachelor	36,8%*	41,7%
	HBO, WO master	48,7%*	7,9%
	PhD or higher	3,4%*	0,6%
	Not educated/unknown	0,3%*	17,1%
Familiarity with Amsterdam	I live there	4,8%	
	I lived there	7,9%	
	Visit once a month	14,6%	
	Visit few times a year	51,6%	
	Been few times	20,6%	
	Never been	0,5%	
Possible to go to Amsterdam by	Own car	61,4%	
	Shared car	22,5%	
	Public Transport	93,9%	
	P+R	51,1%	
	Bicycle	15,1%	

When comparing the age distribution of participants with that of the Dutch population, the first thing that stands out is the number of participants under the age of 35 (57,7%), which is much higher than expected based on the population (42,0%). Due to income differences, the age difference might have resulted in more people choosing to travel by public transport and fewer people choosing to travel by their own car. Possible explanations include that many people in my network are younger than 35 years old and that people in this age group often attend music festival events (Götting, 2021), making it interesting for them to participate in this research.

Another notable difference can be observed among people over the age of 66. The network-related reason also affects this group, as there are few people in my network and among the employees of Sweco who belong to this category. Additionally, the survey was only available online, and people in this age group are often less tech-savvy. Furthermore, this group is not typically a fanatical visitor of mass events, making this research less relevant to them. The average age for festival attendees in Europe was between 21 and 25 in 2016 (Götting, 2021), and for football matches in the Netherlands, it was 45 years old in 2019 (Ruesink, 2019). Because the average age of the visitors of these mass events is lower than the average age in the Netherlands, the average age of the participants should actually be compared with a lower age, which ensures that the differences are smaller.

For the characteristic gender, it is noteworthy that CBS (2022) only included categories for men and women in their distribution. The male-female distribution of participants is almost equal to that of the Dutch population. The differences are not significant.

In the third characteristic, which is the highest level of education completed, there are quite a few significant differences compared to the Dutch population. The participants are generally higher educated than the average Dutch resident. This difference also has to do with how survey responses were collected. Many people within my network and Sweco's employees are highly educated. In addition, it is known that many people who answered 'secondary education' are currently pursuing an HBO or WO bachelor's degree. This gives a somewhat distorted picture of the level of education. Another difference is visible in the category 'not educated/unknown'. At Statistics Netherlands, this is quite a high percentage, while it is a very low percentage among the participants. This is because it was a mandatory question in the survey, which ensured that the education level of all participants was known. In general, it can be said that the participants are more educated than the average Dutch population.

Most people visit Amsterdam several times a year (51,5%), and 12,7% are very well known in the city because these people live or lived there. This is expected to affect the way people travel and prepare for it. The results cannot be compared with those of the Dutch population because these numbers are not known by Statistics Netherlands. The answers to the question about the transport mode options with which people can travel to Amsterdam are partly related to their familiarity with Amsterdam. People who live in Amsterdam or the surrounding area can mostly cycle to the city. If people do not visit Amsterdam often, they do not live in that city, increasing the chance of travelling a greater distance to get there, which is not possible by bicycle. It is striking that the number of people who indicated that they can travel by bicycle to Amsterdam is much higher than those who indicated that they live in Amsterdam. Perhaps not every participant understood the question correctly by answering the question as 'I can travel by bicycle' instead of 'I can travel by bicycle to Amsterdam'.

It is also striking that some people are not willing to travel by certain modes of transport. 61,4% indicated that they can travel to Amsterdam by their own car. In theory, the number of people who can travel to Amsterdam with a shared car should be higher than this number because this includes people who do not own a car but have a driving licence. However, this is not the case (22,5%). The same applies to travelling via a P+R location. People who own a car and can travel by public transport (almost everyone (93,9%)) could, in theory, also use this option, but not everyone is willing to.

Overall, it can be concluded that the participants represent the Dutch population quite well based on gender but not so well based on age and education. However, since this research focused on mass event visitors, the age differences from the population are not such a big problem. The differences from the Dutch population might be caused by how the survey was distributed. Furthermore, not everyone theoretically able to use certain modes of transport actually wants to use these modes of transport.

4.2. Descriptive data set results

Before conducting the analysis models EFA and LCCA, a check was done to identify any striking general results, described in this chapter. The sections are organised by order of the survey questions and cover aspects related to sending information messages. When the events are compared, they are always presented in tables and figures in the same order: music festival, light festival, and football match.

The preferred transport modes for each event were compared, as were the figures from CBS in Section 4.2.1. The various aspects of the information message were discussed next, starting with analysing the moment people prepare for their trip. These values were compared with the moments people want to receive information messages, and again, the values of the events were compared, of which the results can be found in Section 4.2.2. As the second information-related component, the desired message content was analysed in Section 4.2.3, and differences between the events were examined. The third information-related component concerns the channels participants consult to find information, which was analysed in Section 4.2.4. A distinction was made between three phases, before departure, en route, and at the location. Finally, an overview of people's preferred channels for receiving information messages was provided in Section 4.2.5. The main findings were summarised in Section 4.2.6.

4.2.1. Preferred transport mode

The first survey question asked participants about their preferred mode of transport. The options were: own car, shared car, public transport, car to P+R and continuing with public transport, and bicycle. This information is important to be able to send visitors more personalised information. The distribution can be seen in Figure 4.2. An overview of the values, including the frequencies and percentages, can be seen in Appendix D, Table D.1.

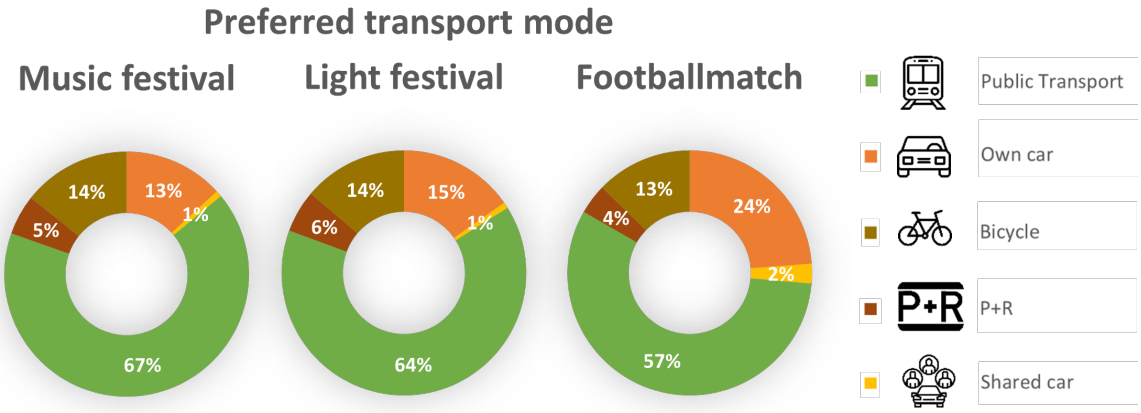

Figure 4.2: Survey results of preferred travel mode per event type

The distribution of travel modes for trips related to entertainment, sports or hobbies, according to the StatLine CBS overviews (CBS, 2022), is presented in Table 4.2. CBS did not include separate categories for shared cars and P+R facilities. Shared cars were assumed to be included in the 'car (driver and passenger)' category. The P+R facility is a combination of car and public transport and cannot be included in one of the categories mentioned above. The exact numbers could not be compared, but in general, it is striking that the share of public transport is much higher in this study, whereas the share of cars is much lower, especially for music festivals. The participants' lower average age could explain this difference compared with the CBS results. Music festivals often involve alcohol consumption, and music festival organisations often promote travelling by public transport and sometimes use extra buses. The alcohol consumption and extra buses might have resulted in a higher share of public transport for music festivals compared with other events.

When the events are compared, it is striking that the distribution of the music festival and the light festival is almost the same. The football match differs slightly from this, but only in the shift from public

transport to using own cars. Bad experiences with public transport, such as on the night of the Kensington concert, which also included a football match at the Arena (Heyblom, 2022), can cause people to switch modes of transport. The people who switched modes of transport showed no clear differences in age or gender compared with the average visitor. Overall, it can be said that public transport is a popular mode of transport in this research, only for the football match some people shifted from public transport to car.

Table 4.2: Travel mode distribution for trips for entertainment, sports or hobbies (CBS, 2022)

Mode of transport	Percentage
Car (driver and passenger)	34,5%
Public Transport	47,9%
Bicycle	3,7%
Walking	2,6%
Other	11,3%
Total	100,0%

4.2.2. Timing

The survey included a question about participants' preparation moments and a question about their preferred time to receive information (communication moment). The differences between the results of these questions are interesting, as they indicate whether participants want to use messages for preparation or only see them as supplements. The timeline used for the questions is not exactly the same, as some categories are not important for certain questions. For example, 'during the trip' is not a preparation moment, and 'travel company prepares the trip' is not a communication moment. However, five categories match, allowing general comparisons between the question results. This can be seen in Figure 4.3, where PREP. represents the preparation moment, and COM. represents the preferred communication moment. The order of events is music festival, light festival, and football match, visualised with icons. The complete data, including frequencies and percentages, can be found in Appendix D, Table D.2.

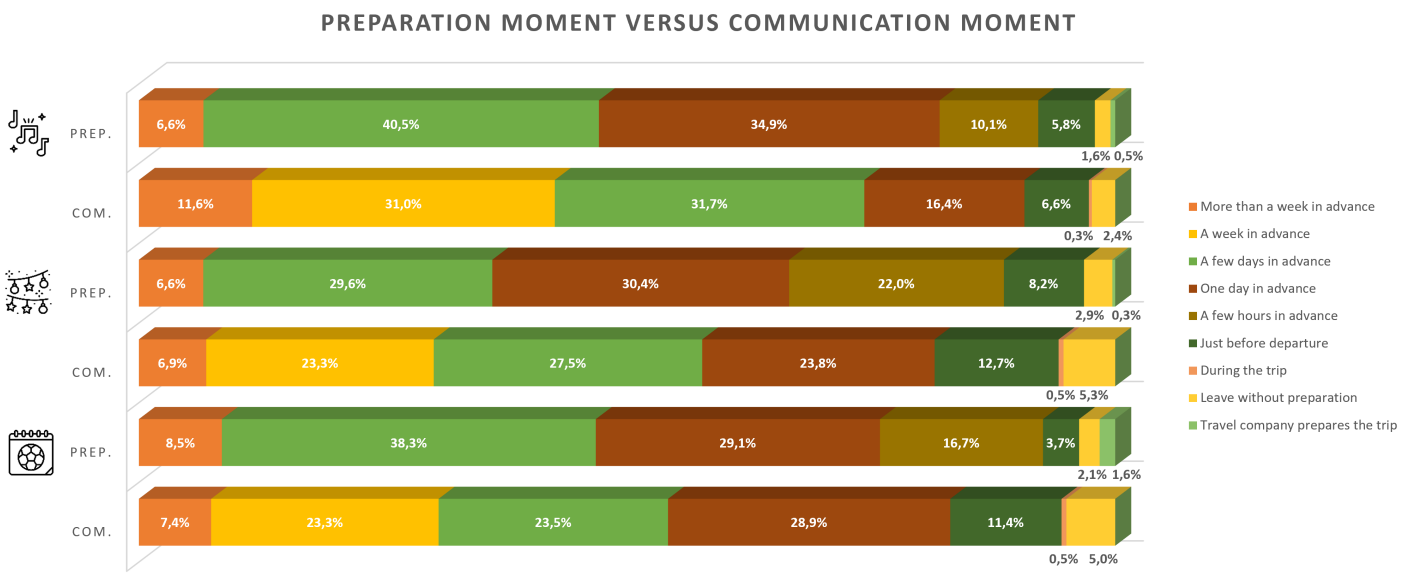

Figure 4.3: Preparation moment compared with preferred communication moment per event type (PREP.= preparation moment and COM.=communication moment)

Figure 4.3 shows that participants generally want to receive messages before preparing their trip, which is usually at least one day in advance (often more). This can be seen in Figure 4.3 because the blocks on the left are relatively larger than the blocks on the right side for COM compared with PREP per event. It shows that messages should be sent further in advance, but not too long in advance, because only 8,6% indicated that they want to receive the information more than a week in advance. Only a few people indicated they do not prepare for their trip and trust their travel company entirely. This suggests that there are many opportunities to inform people during the preparation phase, but that this should not happen too shortly before departure to still be able to influence the choices.

When the different events were compared, it became clear that the preparation moments for the music festival and football match were almost similar, whereas the preparation moment for the light festival occurred a little later. However, this did not correspond fully to the differences in the desired communication moment, where the order was music festival, football match, and then light festival. The order differences are caused by the time difference between the preparation and reception of information, which was smaller for the football match than for the other two events. This desired communication moment order may be related to the frequency of these events (football matches happen almost weekly) and the importance of arriving within a specific time frame, which is not as critical for the light festival.

4.2.3. Content

It is important to provide visitors with the necessary information and avoid bombarding them with additional information to increase the likelihood of them taking action on the information provided. The desired content is illustrated in Figure 4.4, with participants being allowed to select multiple answers. Therefore, the maximum number in each column is 378. The complete data, including frequencies and percentages, can be found in Appendix D, Table D.3.

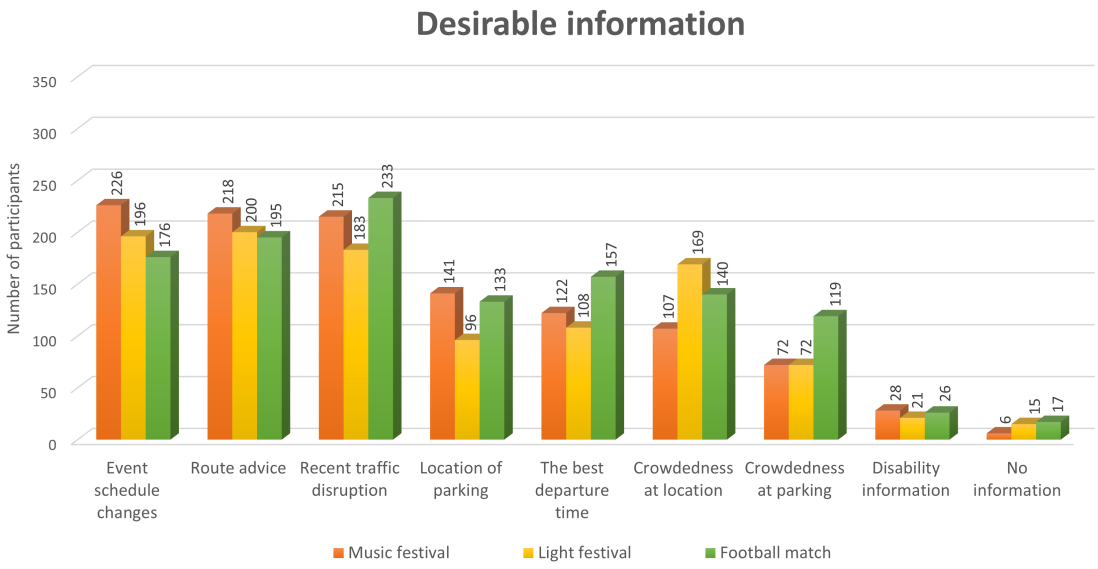

Figure 4.4: Desirable travel or event information per event type

The data does not differ greatly per event, but it is noteworthy that many types of information were selected. On average, over the three events, 1118 types of information were chosen. When viewed per person (by dividing by 378), this means that each participant, on average, selected three types of information. The distribution can be seen in Figure 4.4. Recent traffic disruptions were, on average, mentioned as the most interesting type of information, followed closely by route advice and event schedule changes. These were the most frequently mentioned types of information for all three events. Only a few people indicated they did not want to receive information at all for any of the events. Since people are, on average, interested in three types of information (mostly recent traffic disruptions, route advice, and event schedule changes) and almost everyone wants to receive it, there are many possibilities to inform people with the information they are interested in.

The average number of travellers indicating that they would prefer to travel by car, according to Figure 4.2, is 18%, which corresponds reasonably well with the number of people who want to know more about the crowdedness at the parking ($0,18 * 378 = 68$). It is striking that many more people are interested in the car park location compared with the level of crowding, especially for music festivals. It is not clear from this data why more people want this information. It seems that people who travel by car like to receive information about the parking situation, and some non-car travellers also want information about the parking location.

When the events are compared, it can be seen that information about traffic disruption and the best departure time is less important for the light festival than for other events. A possible explanation is that this event is free for people to join whenever they want. This could also explain why people like to know more about the level of crowding at the location; they can choose a different start time, which reduces crowds and makes parking information less important for this event. Parking information is a frequently mentioned factor for the football match, which is in line with Figure 4.2, which showed that more people prefer to travel by car to this event. Additionally, this event has a fixed start time and, therefore, less flexibility in terms of parking. This limited flexibility also means that when people visit a football match, they are more interested in recent traffic disruption.

4.2.4. Consulted channels

It may be known what kind of information people want to receive and at what time, but how do you deliver that information to them? For this reason, this study included a question asking which channels participants use to access travel and event information. A distinction was made between three moments: before departure, en route, and at the location. The results can be seen in Table 4.3. The participants were allowed to give multiple answers. Therefore, the maximum number in each cell is 378.

Table 4.3: Preparation channels per event type and per moment

		Before departure		**En route**		**At location**	
		Frequency	Percentage	Frequency	Percentage	Frequency	Percentage
Music festival	The organisation's website	303	80,2%	46	12,2%	49	13,0%
	Event organiser emails	268	70,9%	38	10,1%	31	8,2%
	SM channels of the organisation	127	33,6%	47	12,4%	43	11,4%
	SM channels of other organisations	44	11,6%	26	6,9%	19	5,0%
	organisation's app (if available)	95	25,1%	64	16,9%	131	34,7%
	Navigation systems such as Google Maps	207	54,8%	257	68,0%	66	17,5%
	Total sources consulted	1044		478		339	
Light festival	The organisation's website	309	81,7%	48	12,7%	60	15,9%
	Event organiser emails	196	51,9%	22	5,8%	15	4,0%
	SM channels of the organisation	118	31,2%	47	12,4%	40	10,6%
	SM channels of other organisations	47	12,4%	21	5,6%	19	5,0%
	organisation's app (if available)	97	25,7%	51	13,5%	108	28,6%
	Navigation systems such as Google Maps	193	51,1%	231	61,1%	101	26,7%
	Total sources consulted	960		420		343	
Football match	The organisation's website	236	62,4%	35	9,3%	26	6,9%
	Event organiser emails	218	57,7%	36	9,5%	19	5,0%
	SM channels of the organisation	98	25,9%	41	10,8%	34	9,0%
	SM channels of other organisations	39	10,3%	23	6,1%	15	4,0%
	organisation's app (if available)	91	24,1%	59	15,6%	77	20,4%
	Navigation systems such as Google Maps	225	59,5%	227	60,1%	68	18,0%
	Total sources consulted	907		421		239	

The most striking result is that way more channels were selected in the 'before departure' category than for 'en route' and 'at location'. The difference is so big that even if the last two categories are combined, there is still a difference of more than 200 for each event. This means that it is much more difficult to reach people when they are en route or already at the location than before departure. This is not very consistent with the literature, in which people are often informed when they are en route. The results of this study indicate that this is not the best moment. People can be reached most easily before departure, making this the most important opportunity to inform people. Therefore, after identifying the

most striking results, this study has decided to focus on the situation before departure, also known as the preparation moment.

The differences between the events were analysed per moment using a Chi-Square test of independence using $\alpha = 0,05$. The results can be seen in Table 4.4, with M = music festival, L = light festival, and F = football match. It can be seen that the results en route do not differ significantly (p-value above 0,05), while the results before departure and at location do differ significantly. En route, people are most consistent in using the same information channels, especially navigation systems. A possible explanation is that people en route mainly focus on their route, but they also look up other information before departure and at the location.

Table 4.4: P-value results of Chi-Square test of independence where the events are compared per moment

	Before departure	**En route**	**At location**
p-value: M compared with L	0,027	0,435	0,004
p-value: M compared with F	0,023	0,756	0,023
p-value: L compared with F	0,004	0,250	0,025

Before departure, it was striking that many sources were consulted, an average of 2,6 per person ($(1044 + 960 + 907)/3/378 = 2,6$), with the organisation's website being the most frequently mentioned for all three events. Shortly after, emails from the organisation and navigation systems were the most mentioned. These three categories (almost) cover three-quarters of the total sources consulted and are, thus, important channels for the organisation to focus on. Social media channels from other organisations were mentioned the least. This is a good result because this is a difficult channel through which to inform.

En route, most participants seem to use navigation systems to access information, which scored very high compared with other channels. Since navigation systems provide information about the route, little other information is being looked up en route. This is a channel through which it is currently hard to inform within CROMAS, which means that for now, the 'en route' phase is less relevant for CROMAS.

At the location, participants indicated they mainly use the event app and navigation systems, but not very often. The website was also ranked well, especially for the music festival. This is expected to be related to event information, such as the timetable of artists being available on the app and website. However, these scores are very low compared with the 'before departure' category. The somewhat higher score for navigation systems at the light festival might be related to the fact that people are allowed to determine their route at this event and have to keep an eye on it. The results show that people do not really focus on their return trip when they are at the location, and suggest that it is better to send information about the return trip before departure. It is difficult to get this information to visitors through these channels after departure, so other ways to reach visitors after departure should be looked at if important information needs to be shared.

In the survey, participants mentioned public transport planning apps, like 9292 and NS, a lot as an additional answer (61, 50, 46 per event). Some participants may have classified this under navigation systems, but many did not. Since not everyone may have done this, the values do not provide a complete picture of the numbers. It is good to include this answer option if the question is asked again. However, these channels are less important for CROMAS because they cannot obtain information through them.

Since this study focused on pre-departure data, the distributions for this category have been examined more closely and are visualised in Figure 4.5. It is striking that emails seem much more important when preparing for a trip to a music festival compared with the other two events. This is not compensated for by another category because the total number for the music festival is higher than for the other events. The low score for the website at the football match can possibly be explained by the fact that the event takes place almost weekly, and the visitors already know the general information and only look for changes.

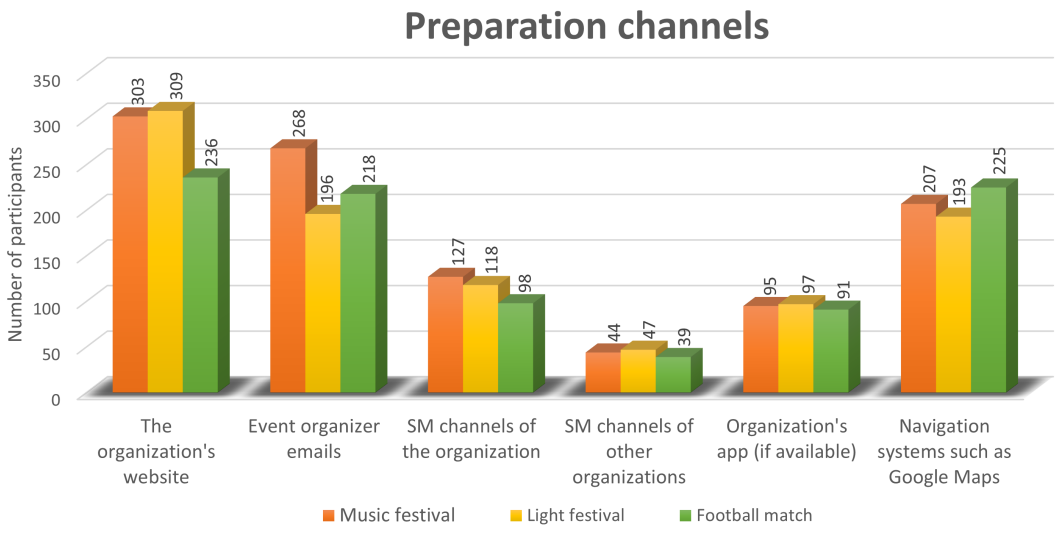

Figure 4.5: Channels consulted before departure

4.2.5. Communication channels

To specify further how to reach visitors best, the participants were asked which channels they prefer to receive information through. A list of social media channels was provided, along with event apps, email, and SMS. These channels were selected because they allow messages to be sent easily by CROMAS (now or soon). Participants were allowed to give multiple answers, so the maximum number in each column is 378. The results are presented in Figure 4.6. The full data, including frequencies and percentages, can be found in Appendix D, Table D.4.

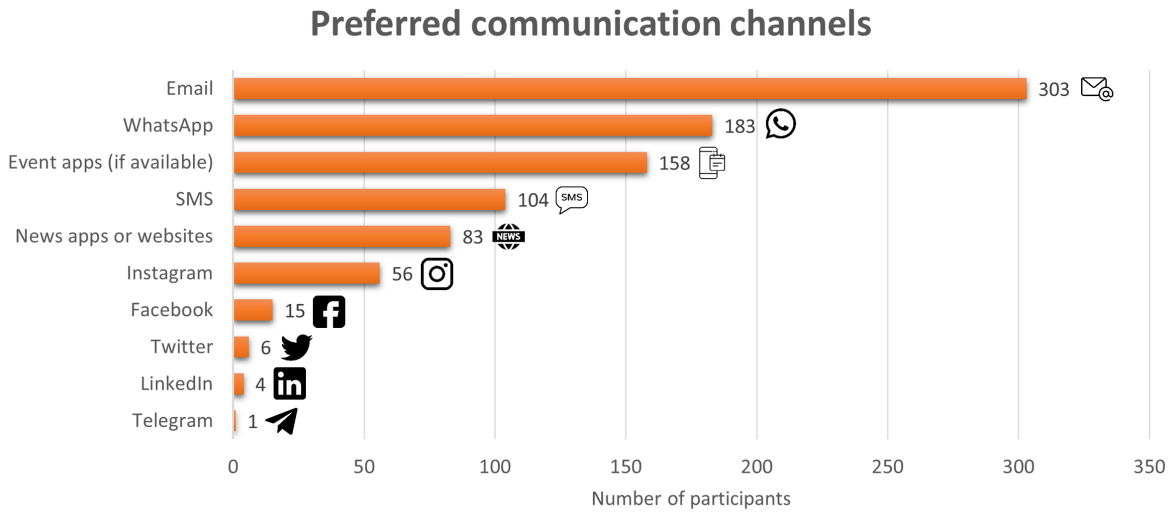

Figure 4.6: Preferred information communication channels

The most notable value is for email, which is considerably higher than the other values. Other high values can be seen for WhatsApp, event apps, and SMS. These sources have in common that they are originally intended for sending messages. Channels primarily used for entertainment by scrolling through posts like Facebook, Instagram, and Twitter were mentioned less frequently. Popular social media channels do not seem to be the best way for participants to receive important information, even though event organisations often focus on these channels. This may be due to the different ways in which the channels present information and the extent to which the information is easy to retrieve.

4.2.6. Main findings

The most striking results from the previous sections, which in some cases led to important analysis choices, have been combined below:

- The results reveal a lot of overlap between the different events.

  → **It was decided to apply multiple factor analyses to investigate whether the events differ so little that they can be combined into a factor.**

- A considerable number of participants prefer public transport to travel to these events.
- Participants for all three events desire to receive messages at least one day in advance, and preferably earlier (maximum one week in advance), before preparing for their trip. People want to be informed, but this should not happen too shortly before departure to still be able to influence the choices.
- Very few participants indicated that their travel company prepares the trip, which means that it is important that everyone receives the information.
- People are, on average interested in receiving three types of information (mostly recent traffic disruptions, route advice, and event schedule changes), and almost everyone wants to receive it. So, there are many possibilities to inform people with the information they are interested in.
- People look up more information before departure than en route or at the location. They do not really seem to focus on their return trip when they are at the location. According to the results of this study, it is important to provide visitors with all information (including the return trip) before departure.

  → **This study focuses on pre-departure data.**

- The website, emails, and navigation systems are the most commonly used preparation channels among participants and are therefore important for the organisation to focus on.
- Participants clearly prefer receiving information via email and do not prefer receiving import information via social media.

4.3. Exploratory Factor Analysis

The steps described in Section 3.2 and visualised in Figure 3.3 were used to perform the Exploratory Factor Analysis. As stated in Section 4.2.6, it was examined whether certain data correlated enough to be combined. The EFA was performed for the moment, content, and channels, explained in the coming sections. For all analyses, PCA was used as the factor extraction method (step 2), and the orthogonal rotation with the Varimax technique was used (step 4). The abbreviations M (music festival), L (light festival) and F (football match) were used to indicate the events. The mentioned survey questions can be found in Appendix C. The results of the EFA were used as input for the Latent Class Cluster Model.

4.3.1. Moment

The factor analysis of the moment includes the answers to six survey questions. Three sets of answers relate to when participants prepare for their trip (one for each event), and the other three relate to when they want to receive travel information (one for each event). This is referred to by the abbreviations 'PREP' and 'COM'.

First, it was checked whether the data were suitable for factor analysis. The results showed a KMO value of 0,712 and a significant Bartlett's Test ($p<0,001$). Since the KMO value was higher than 0,5 and Bartlett's test showed significance, the data were suitable for factor analysis. Furthermore, all the extractions of communalities were larger than 0,3. When determining the number of factors, it emerged that one factor would be the most appropriate, meaning that all factors correlate. Therefore, the correlation matrix was checked, shown in Table 4.6. All correlations are significant at the 0,01 level (2-tailed).

So, all variables can be combined based on this criterion. This result is not unexpected because Figure 4.3 already showed a lot of overlap between the different preparation and communication moments. The last step included analysing Cronbach's Alpha value, which was equal to 0,834. This value is larger than 0,8 (very good), meaning that the items are closely related and can be combined in one factor.

Table 4.6: Correlation matrix moment, where PREP = preparation moment and COM = communication moment, for the events M = Music festival, L = Light festival, F = Football match

	M PREP	M COM	L PREP	L COM	F PREP	F COM
M PREP	1	0,479	0,596	0,312	0,430	0,337
M COM	0,479	1	0,385	0,619	0,387	0,549
L PREP	0,596	0,385	1	0,529	0,474	0,334
L COM	0,312	0,619	0,529	1	0,361	0,459
F PREP	0,430	0,387	0,474	0,361	1	0,635
F COM	0,337	0,549	0,334	0,459	0,635	1

4.3.2. Content

During the factor analysis about the content, the answers to the questions about which content the participants would like to receive were included per event. It was decided not to include the answer option 'no information' in the analysis because this answer can be deduced from the other answers. This resulted in 24 items (eight for each event). The label names refer to the answer options of questions 4, 9, and 14 of the survey (Appendix C).

The suitability was checked by looking at the KMO value of 0,718 and the significance of Bartlett's Test (p=0,000). Since the KMO value is higher than 0,5 and Bartlett's test showed significance, the data was suitable for factor analysis. Furthermore, all extractions of communalities were larger than 0,3.

Second, the criteria which were used to determine the number of factors (Kaiser rule, scree plot, and parallel analysis) were checked, resulting in 7 factors, visualised in Table 4.7. Each factor satisfies the fact that each factor must contain at least two to three variables. The reliability of the factors was tested by calculating Cronbach's Alpha values, presented in Table 4.8. All values are larger than 0,6 (accepted), meaning that the items are closely related and can be combined in one factor. No variables had to be deleted.

It can be seen that the factors mainly combine the events and no other information types apart from the parking-related information. This ensures that the combinations can be theoretically substantiated. The events do not appear to differ enough to result in different factors. Furthermore, the interest in different types of information appears to differ too much to be combined. The only exception to this can be seen in the parking location and parking crowdedness. Of all types of information, these two are most related, which makes the combination theoretically logical, as they both have to do with the parking situation. This distribution aligns with the expectations because the information content differs a lot, except for the parking information. It was previously seen that the different events show overlapping results, so the combination of the events in one factor is not surprising.

Table 4.7: Rotated Component Matrix of the information content, for the events M = Music festival, L = Light festival, F = Football match

Item	1	2	3	4	5	6	7
Informatie_parkeren							
parkeerlocatie_F	0,756						
parkeerlocatie_L	0,75						
parkeerlocatie_M	0,743						
parkeerdrukte_F	0,743						
parkeerdrukte_L	0,67						
parkeerdrukte_M	0,633						
Informatie_verkeershinder							
verkeershinder_L		0,86					
verkeershinder_M		0,848					
verkeershinder_F		0,732					
Informatie_beperking							
beperking_L			0,874				
beperking_M			0,827				
beperking_F			0,798				
Informatie_routeadvies							
routeadvies_M				0,807			
routeadvies_L				0,802			
routeadvies_F				0,738			
Informatie_bestevertrektijd							
bestevertrektijd_M					0,795		
bestevertrektijd_L					0,769		
bestevertrektijd_F					0,692		
Informatie_evenementschema							
evenementschema_L						0,809	
evenementschema_M						0,802	
evenementschema_F						0,669	
Informatie_drukteterplaatse							
drukteterplaatse_M							0,766
drukteterplaatse_L							0,721
drukteterplaatse_F							0,69

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization. Rotation converged in 6 iterations.

Table 4.8: Cronbach's Alpha of each factor

Factor	Cronbach's Alpha
1: Informatie_parkeren	0,812
2: Informatie_verkeershinder	0,797
3: Informatie_beperking	0,791
4: Informatie_routeadvies	0,732
5: Informatie_bestevertrektijd	0,702
6: Informatie_evenementschema	0,703
7: Informatie_drukteterplaatse	0,612

4.3.3. Channels

the factor analysis of the channels included 18 items (six for each event), which were the answers to the questions about which channels the participants use to look up the event and travel information (Q3, Q8, and Q13).

The suitability of the data for factor analysis was analysed by checking the KMO value (0,797) and the significance of Bartlett's test (p=0,000). Since the KMO value is higher than 0,5 and Bartlett's test

showed significance, the data was suitable for factor analysis. Furthermore, all the extractions of communalities were larger than 0,3.

The Kaiser rule, scree plot and parallel analysis criteria resulted in 4 factors, which can be seen in Table 4.9. Each factor consists of three or six variables, which is more than the minimum of two to three. Cronbach's Alpha values represent the reliability of the factors, and can be seen in Table 4.10 for all four factors. All values are larger than 0,6 (accepted), meaning that the items are closely related and can be combined in one factor. No variables had to be deleted.

According to Table 4.10, the factors are all reliable, but the results must also be theoretically substantiated. Again, it can be seen that the factors mainly combine the events. Furthermore, the social media related channels (SM) were combined. This was an expected combination since the same channels were used, viewing different profiles. Furthermore, the mail and website have been merged. Table 4.3 already showed that these are two channels that the participants often use to prepare their trip. Both channels are somewhat more conservative sources than social media and apps, which ensures that this combination is not very surprising. In general, the factors show a lot of overlap between the different events, and a few sources overlap.

Table 4.9: Rotated Component Matrix of the trip preparation channels, for the events M = Music festival, L = Light festival, F = Football match

Item	1	2	3	4
Voorbereidbron_SocialMedia				
SMderden_L	0,772			
SMderden_F	0,755			
SMderden_M	0,706			
SMorga_F	0,688			
SMorga_M	0,673			
SMorga_L	0,636			
Voorbereidbron_WebsiteMail				
website_L		0,724		
website_M		0,691		
mail_M		0,624		
mail_F		0,592		
mail_L		0,584		
website_F		0,508		
Voorbereidbron_Navigatie				
navigatie_L			0,872	
navigatie_M			0,844	
navigatie_F			0,82	
Voorbereidbron_App				
app_L				0,782
app_M				0,762
app_F				0,743

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
Rotation converged in 6 iterations.

Table 4.10: Cronbach's Alpha of each factor

Factor	Cronbach's Alpha
1: Voorbereidbron_SocialMedia	0,821
2: Voorbereidbron_WebsiteMail	0,735
3: Voorbereidbron_Navigatie	0,811
4: Voorbereidbron_App	0,754

4.4. Latent Class Cluster Analysis

The Latent Class Cluster Analysis groups individuals with comparable behavioural patterns regarding their attitude towards information messages. The steps explained in Section 3.3 were executed for the model shown in Figure 3.5. The EFA reduced the number of indicators used as input for the LCCA. To calculate the factor scores, the results of the multiplication of the variable answer and the factor coefficients were summed up, and divided by the number of variables in that factor. So if a factor consists of three variables, each variable is multiplied by the factor coefficient, and then the three results are added and divided by three. This was done for each factor.

First, the measurement model was estimated with all indicators. For this model, this means 12 indicators (1 moment, 7 information, and 4 channels). As explained in Section 3.3, the number of clusters was determined by looking at the BIC, AIC, and BVR values. However, when all indicators were included in the model, the results showed very high BIC values and many significant BVR values, where low BIC values and insignificant BVR values are preferred. The best model fit resulted in a BIC value of 13292,41 and 12 significant BVR values. Many significant BVR values indicate many significant relations between indicators, which is not desirable in this search for independent clusters. Therefore, the number of indicators had to be reduced.

Whether indicators were included in the model or not was based on theory. The number was reduced until a suitable model was found. In Figure 3.5, it can be seen that the communication channels occur as a covariate and as an indicator. The covariate is about the moment the participants prefer receiving information, and the indicator is about the channels they use to prepare for their trip. Since this research focuses on sending information, the data about preferred receive channels is more relevant. Therefore the indicators about the channels were removed from the model. Additionally, one of the content factors is about disability information. Since only a few people indicated that they would like to receive this information (as mentioned in Table 4.3), it is unlikely that this information will be disseminated via information messages, making it less interesting to include this indicator in the model. In the end, a new model was estimated with 7 indicators (1 moment and 6 information), resulting in the model fit statistics visualised in Table 4.11.

Table 4.11: Model fit statistics for the LCCA,
where LL = Log-Likelihood, BIC = Bayesian Information Criteria, AIC = Akaike Information Criterion, Npar = Number of parameters, Δ%BIC = percentage difference in BIC, and #- Sig BVR = Number of significant Bivariate Residuals

#-Clusters	LL	BIC(LL)	AIC(LL)	Npar	#- Sig BVR
1-Cluster	-3654,52	7469,28	7363,04	27	14
2-Cluster	-3587,88	7383,48	7245,76	35	5
3-Cluster	-3564,08	7383,37	7214,17	43	0
4-Cluster	-3548,54	7399,77	7199,09	51	0
5-Cluster	-3537,45	7425,06	7192,90	59	0
6-Cluster	-3528,38	7454,40	7190,77	67	0
7-Cluster	-3516,48	7478,08	7182,96	75	0
8-Cluster	-3504,78	7502,16	7175,57	83	0
9-Cluster	-3496,08	7532,24	7174,17	91	0
10-Cluster	-3489,23	7566,01	7176,46	99	0

From the fourth cluster onwards, it can be seen that the BIC increases (a low value is desired), and the AIC improvements become smaller. Furthermore, the models with one and two clusters showed significant BVR values (Wald value > 3,84), whereas the other models with three or more clusters did not show significant BVR values. This means that there is no significant relation between indicators. After analysing the criteria, the 3-cluster model was the most suitable model for this research.

The model ensured that, based on the indicators included, it is known to what extent it is likely that a participant belongs to each cluster. The covariates were added to the measurement model to

create the structural model, to find out the characteristics of the people belonging to each cluster. 13 covariates were added to the model by applying backward elimination until all remaining covariates were significant (Wald values higher than 3,84 or p-value below 0,05). This was done by adding all 22 covariates, re-estimating the model 22 times (each time without a single covariate), and setting the covariate with the smallest effect on the model fit as inactive. The model was then re-estimated with 21 covariates, and the steps were repeated until all remaining covariates showed significance, resulting in 13 active covariates and 9 inactive covariates. Finally, the entropy R-squared value was examined to see how accurately the model calculated the clusters in combination with the covariates. The entropy R-squared value was 0,9381, which is larger than 0,8, meaning, according to Clark and Muthén (2009), ”it was most likely class membership was the best performing method in terms of recovering the true value used in the simulation study.”

The cluster profiles are explained by numbers in Table 4.12, including the sample shares, indicators, active covariates, and inactive covariates. The extent to which each cluster is interested in receiving certain information is shown in Figure 4.7. The values indicate how often someone indicated to be interested in that type of information. The participants could answer this question with yes or no for all three events. So this value can be equal to 0 (0/1), 1 (0,33/1), 2 (0,66/1), or 3 (1/1) times. Here, the first differences between the clusters are already visible. The cluster interpretations are given below the table and the preferences are visualised in Figure 4.8.

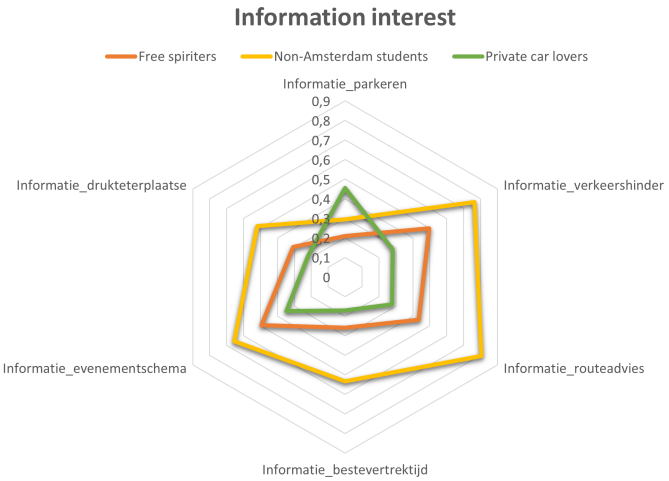

Figure 4.7: Interest in certain information content per cluster, frequency on the range 0 (never) to 1 (always)

Table 4.12: Cluster profiles, including indicators, active covariates, and inactive covariates

Cluster name		**Free spiriters**	**Non-Amsterdam students**	**Private car lovers**
Cluster size		49%	35%	16%
INDICATORS (MEAN)				
	Informatie_parkeren	0,2099	0,2949	0,4548
	Informatie_verkeershinder	0,4972	0,7645	0,2814
	Informatie_routeadvies	0,4356	0,8062	0,2785
	Informatie_bestevertrektijd	0,259	0,5338	0,1703
	Informatie_evenementschema	0,494	0,656	0,3473
	Informatie_drukteterplaatse	0,306	0,5194	0,218
	Moment	3,3283	2,6747	3,5211
ACTIVE COVARIATES				
Age	15-25	29,2%	47%	23%
	26-35	22,2%	24%	25%
	36-45	8,1%	2%	15%
	46-55	17,3%	16%	23%
	56-65	19,5%	10%	13%
	66+	2,7%	2%	2%
	Mean	39	33	39

Gender	Male	62%	34%	61%
	Female	38%	65%	36%
	Other	0%	0%	3%
	Blank	0%	1%	0%
Familiarity with Amsterdam	I live there	10%	0%	3%
	I lived there	11%	2%	10%
	Visit once a month	18%	7%	18%
	Visit few times a year	49%	59%	44%
	Been few times	12%	32%	23%
	Never been	1%	0%	2%
Possible to go to Amsterdam by	Own car	57%	53%	95%
	Shared car	12%	33%	30%
	Public Transport	95%	99%	80%
	Bicycle	20%	5%	23%
Prefer receiving information via	WhatsApp	44%	58%	42%
	Instagram	5%	28%	17%
	News	14%	32%	25%
	SMS	26%	37%	9%
Preferred transport mode M	Own car	4%	6%	57%
	Shared car	0%	1%	3%
	PT	64%	85%	35%
	P+R	10%	1%	2%
	Bicycle	23%	8%	3%
Preferred transport mode L	Own car	2%	7%	75%
	Shared car	1%	2%	0%
	PT	67%	83%	19%
	P+R	10%	1%	2%
	Bicycle	21%	8%	5%
INACTIVE COVARIATES				
Education	Primary education	0%	1%	0%
	Secondary education	9%	17%	2%
	MBO, HBO, WO bachelor	35%	39%	38%
	HBO, WO master	51%	42%	57%
	PhD or higher	5%	2%	3%
	Not educated/unknown	0%	1%	0%
Possible to go to Amsterdam by	P+R	44%	55%	65%
Prefer receiving information via	Facebook	2%	5%	10%
	LinkedIn	1%	0%	5%
	Twitter	2%	1%	3%
	Telegram	1%	0%	0%
	Event Apps	37%	52%	33%
	Email	80%	86%	68%
Preferred transport mode F	Own car	15%	20%	59%
	Shared car	1%	6%	2%
	PT	61%	63%	30%
	P+R	7%	1%	0%
	Bicycle	16%	9%	10%

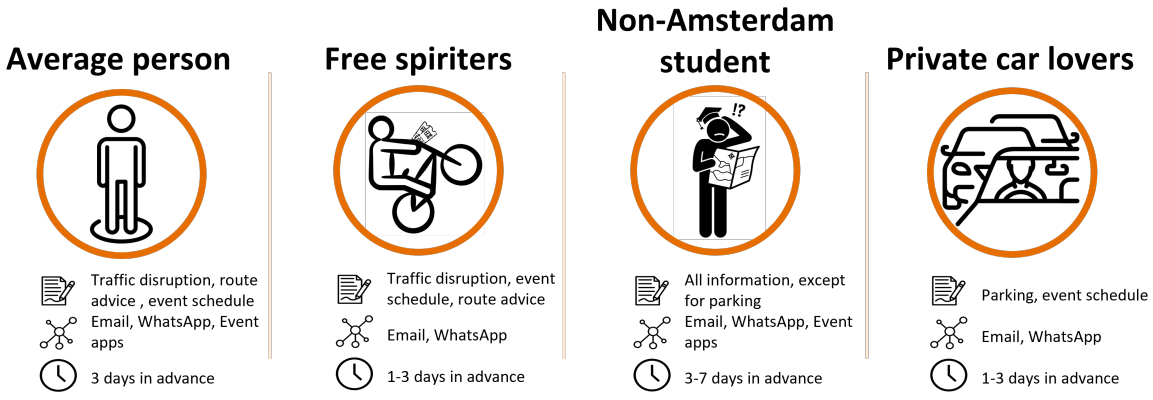

Figure 4.8: Information provision preferences per cluster

4.4.1. Free spiriters

Almost half of the participants (49%) belong to the first cluster. The typical *free spiriter* has several transport mode options at its disposal but considers which transport mode they prefer to travel with each time. Many of the *free spiriters* live in Amsterdam or visit Amsterdam regularly. As a result, they have learned that it is not easy to move around by car in Amsterdam. This has ensured that the car is not the first choice, even though many *free spiriters* have one available (57%). They are more outspoken about the channels through which they want to receive information. They do not like the new social media platforms, such as Instagram, Facebook, LinkedIn, Twitter, and Telegram. The preference is clear: communication via email (80%) and, otherwise, WhatsApp messages are accepted (44%). This group is generally highly educated and mainly consists of men (62%).

Receiving information does not seem necessary, as can be seen in Figure 4.7, especially car information (0,21/1). Probably, because they have no intention of travelling by car anyway. Travelling with less information appears to require more experience. They know how crowded it can be in Amsterdam and might therefore be less interested in information about crowdedness (0,31/1). If information is sent, the messages are preferred to contain less self-evident information, such as traffic disruption (0,50/1) or events schedule changes (0,49/1), not too long in advance.

4.4.2. Non-Amsterdam students

The second cluster includes 35% of the participants and is referred to as *non-Amsterdam students* since this group only visits Amsterdam a few times a year or less. They are unfamiliar with the city. Additionally, their average age is much lower than that of the other clusters, 33 compared to 39. Almost half of this group (47%) is between 15 and 25 years old. This group scores below the average on 'highest completed education' and is younger, which suggests that they are still studying. The lower average age also affects their transportation options. The percentage of car owners in this group is lower than in the other clusters (53%), and almost everyone (99%) can travel to Amsterdam by public transport. Therefore, they prefer this transport mode when travelling to events in Amsterdam. The fact that they are unfamiliar with Amsterdam is also reflected in the low percentage of people who can travel to Amsterdam by bicycle (5%). *Non-Amsterdam students* also like to use their phones. Their favourite channel through which they receive information is email (86%), but they also indicated to be happy to receive information from various popular social media channels. They indicated this much more often than the other clusters. It is interesting to note that the somewhat older form of communication, SMS, is also seen as a nice communication channel (37%), but the best days of Facebook (5%) and Twitter (1%) are already over. Notably, this group consists mainly of women (65%), even though part of this group is expected to study in Delft, where there are more male students.

As shown in Figure 4.7, *non-Amsterdam students* crave information. Since they almost always travel by public transport, they find parking information to be the least interesting (0,29/5). However, they are happy to receive all other information, especially messages about route advice (0,81/1) and traffic disruption (0,76/1), preferably at least a few days in advance.

4.4.3. Private car lovers

The *private car lovers* comprise 16% of the participants, mainly men (61%). It is evident that they enjoy driving their car. 95% of them own a car and drive it everywhere, even though this group has previously visited Amsterdam and is aware that it is not always the most convenient mode of transportation in this city. Since they possess a car, they prefer to drive it. Theoretically, they could also use a shared car, but the majority (70%) stated this is not feasible. If an alternative mode of transport is necessary, they prefer public transport. Purchasing a car requires money, which is reflected in their educational level. This group is highly educated and has been working for a while, as indicated by the average age of 39.

Their desired information also demonstrates that *private car lovers* are creatures of habit. They prefer to be free of information, except when it is related to their car, of course (0,45/1). You do not need to send them additional information, which was clearly indicated by their preferred communication channels. They are not particularly enthusiastic about any specific channel, but if they have to choose, they prefer email (68%).

4.5. Behavioural intention

Information messages were shown to the participants to gain an initial understanding of the extent to which individuals behave according to their specified preferences. Each cluster has unique preferences regarding the content and timing of information they receive. The clusters were determined based on the probability of each person belonging to a specific cluster. To compare information preferences with follow-up behaviour, participants were assigned to a cluster by 'hard partitioning', meaning each participant is assigned to the cluster they most likely belong to.

Information messages were used to express the behavioural intention of the participants in numerical values. The participants indicated at five different messages how likely they would act on this information. The messages used for this can be found in Section 3.1.1. Figure 4.9 illustrates how each cluster rated the messages, where the messages are indicated by their topics. The numerical values and percentages can be found in Appendix D, Table D.5. Table 4.13 includes the cluster size, mean score, and standard deviation of the ranking scores for each cluster. Based on this, the differences between the clusters were statistically tested with a two-tailed test. For the degrees of freedom 315, 194, and 241, the significant t-value with a confidence level of 0,95 is 1,984, meaning that the ranking scores differ significantly if a t-value is above this value. The larger the t-score, the larger the difference between groups. The smaller the t-score, the more similarities there are between groups. The asterisk sign in the table indicates significant results. The results are discussed below.

Table 4.13: Two-tailed test with alpha = 0,05, significant values indicated by *

		Departure time	Route	Mode of transport	Parking availability	Public transport timetable
Free spiriters	N	182	182	182	182	182
	Mean	3,32	3,97	3,70	3,40	3,19
	Std. Deviation	0,897	1,17	1,31	1,282	1,097
Non- Amsterdam students	N	135	135	135	135	135
	Mean	3,28	4,14	3,78	3,62	3,33
	Std. Deviation	0,903	1,121	1,262	1,281	1,078
Private car lovers	N	61	61	61	61	61
	Mean	3,03	4,07	3,48	3,44	2,82
	Std. Deviation	1,125	1,031	1,206	1,259	1,148
Free spiriters vs	|t|	0,391	1,299	0,545	1,509	1,130
Non- Amsterdam students	df	315	315	315	315	315
Non-Amsterdam students vs	|t|	1,655	0,414	1,558	0,913	2,998*
Private car lovers	df	194	194	194	194	194
Free spiriters vs	|t|	2,041*	0,593	1,155	0,211	2,249*
Private car lovers	df	241	241	241	241	241

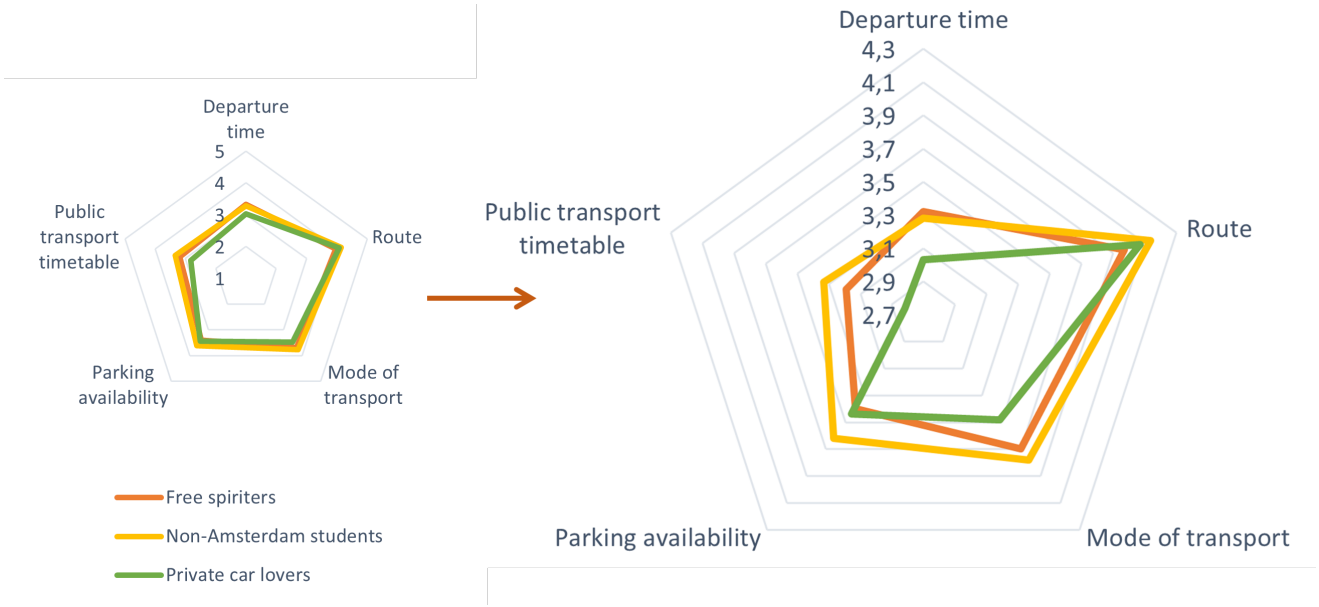

Figure 4.9: Message rankings per cluster, scale ranges from 1 (absolutely not) to 5 (definitely)

Looking at the results in the left diagram, there is a lot of overlap between the clusters. The table also reflects this because only 3 of the 15 t-scores are significant. Therefore, the second diagram was created with an adjusted range, which shows the differences more clearly. Despite this adjustment, the answers are still quite similar. The clusters showed clear differences in information preferences, which was expected to be reflected in their follow-up behaviour. However, this difference did not become clear through these results. For instance, the clusters that do not travel by car frequently (*free spiriters* and *non-Amsterdam students*) still indicated that they would act on the message about parking information. It appears that the participants did not rank the messages with the idea of travelling by a particular mode of transportation or based on their 'cluster status' but rather judged the message based on how they would have responded if they were in that situation.

One striking observation is that the message about route advice was the most popular message among the participants, whereas the messages about departure time and public transport timetables were far less popular. Whether this disparity is solely due to the subject matter or whether there were other factors affecting the results, such as content including not very interesting advice or message wording, needs further investigation. The message about the public transport timetable is assessed significantly differently by the *private car lovers* than by the *free spiriters* and the *non-Amsterdam students*, here the preferred modality of the *private car lovers* seem to have influenced the results.

The biggest differences are between the 'not-so-interested' *private car lovers* and 'very interested' *non-Amsterdam students*. The *private car lovers* previously indicated that they were not particularly interested in receiving information, which was confirmed by these results, even when the information was related to parking availability. The *non-Amsterdam students* expressed interest in a lot of information. The results of figure 4.9 align with this and show that this cluster is most likely to act on information. However, for most messages, their ratings were only slightly higher than the average value of 3, indicating that it is not certain that they would all act on the information. Since they indicated so much interest in different types of information, these results do not match the expectations. Firm conclusions cannot be drawn from these findings because various other factors may be related to this. For example, the participants may not have judged the messages based on their 'cluster status'. Additionally, the wording of the messages or the extent to which the participants can relate to the situation can influence factors. It is also possible that the clusters would have shown larger differences in behavioural intention if the clusters were created based on other factors. The differences between the clusters are expected to be more significant than the results show now. Nevertheless, the results show that the *non-Amsterdam students* are most interested in information and the most impressionable cluster.

5

Conclusions and recommendations

This chapter is the final chapter, in which the conclusions of this research are presented. In addition, the scientific and practical implications are discussed, followed by the limitations. The chapter ends with recommendations for follow-up research based on the conclusion and limitations.

5.1. Conclusion

This research aimed to identify the preferences of different types of visitors towards receiving information and the extent to which visitors with different information profiles act on the information messages they receive. The information is communicated via messages that visitors receive on their mobile phones. The information profile contains a visitor's preference regarding the content, communication channel, and communication timing. In addition, a distinction was made between travelling to three different mass events in Amsterdam, each intended for entertainment. To find the answer to the research question, a quantitative study was conducted on people's preferences regarding the different aspects associated with receiving information messages by distributing a survey (N=378). The results were analysed with an Exploratory Factor Analysis (EFA), a Latent Class Cluster Analysis (LCCA), and a comparison of the behavioural intentions.

The results revealed, in general, that most communication channels are consulted before departure (more than twice as often than en route or at the location). It is much more difficult to reach people en route or at the location because they are not actively looking for information. Therefore, sharing all possible information, including information about the return trip, with the visitors before they leave is important. Everyone must receive this information because hardly anyone (0,8%) lets their travel company prepare their trip. What also stood out is that even though event organisations increasingly focus on social media, these are not the channels through which visitors prefer to receive information. Email is by far the most popular channel to receive information (indicated by 80%). It was remarkable that these results and the results of the information profiles showed few striking differences between the events. There are differences, but the three trip purposes, music festival, light festival, and football match, appeared to influence the preferences to a limited extent. These little differences between the events were also revealed after performing the EFA, which ensured that the different events were merged in factors before the LCCA.

This study distinguished clusters based on their information reception preferences by performing a LCCA. All clusters deviate to some extent from the preferences of the average participant, referred to as the average person in this study. The average person is 37 years old, highly educated (52% HBO, WO master or higher), visits Amsterdam several times a year, and prefers to travel to these events by public transport (63%). The average person prefers to receive information messages before trip preparation. This means one day in advance or more (according to 93% of the participants) but with a maximum of one week in advance. They prefer messages containing information about recent traffic disruptions (56%), route advice (54%), and/or event schedule changes (53%). This information is preferably communicated via email (80%) and otherwise through WhatsApp messages (48%).

The LCCA resulted in three clusters that represent different preferences: 1) *free spiriters* (49% of the participants), 2) *non-Amsterdam students* (35% of the participants), and 3) *private car lovers* (16% of the participants). Each cluster has its preferences and is best approached in a different way. The LCCA clarified the clusters, which are described by their most notable traits and preferences as follows:

- The *free spiriters* change mode of transport, where they usually opt for public transport (64%) or bicycle (20%), and mostly not the car (7%). They are familiar with Amsterdam because they live(d) there or visit it regularly. In addition, they are mainly men (62%) who are highly educated (56% HBO, WO master or higher). They would like to receive information messages about recent traffic disruptions, event schedule changes, and/or route advice by email between 1 and 3 days in advance.
- The *non-Amsterdam students* are unfamiliar with Amsterdam and almost young compared to the other clusters. Half of the group is between 15 and 25 years old. They clearly prefer public transport when travelling to Amsterdam (77%). The *non-Amsterdam students* are very interested in receiving information and less critical about the channel through which they receive it. They would like to receive messages about all information types except parking information. Messages about route advice are most wanted, preferably received by email between 3 and 7 days in advance. This cluster is the most impressionable cluster.
- The *private car lovers* almost all own a car (95%) and clearly prefer to use it to drive to the events compared with the other clusters (64% compared with 7% and 11%), since they are creatures of habit. This cluster consists mainly of men (61%) who regularly visit Amsterdam. They would like to receive information messages about parking and/or event schedule changes by email between 1 and 3 days in advance. This cluster is the least impressionable cluster.

The extent to which these clusters differ in behavioural intention is small in this study. The more interested *non-Amsterdam students* show a higher likelihood score in all categories than the less interested *private car lovers*. Still, the differences are only significant for the departure time and public transport timetable. The participants did not seem to have rated the messages from their 'cluster status' but from a perspective where the message matches their travel choices. The message about route advice was the most popular among the clusters. No firm conclusions can be drawn from the results. These results give a first indication of the behavioural intention.

5.2. Implications

No previous studies were found using the LCCA method to investigate information message preferences or how to influence travel behaviour. Travellers were often pre-assigned to a group based on their mode of transport or trip purpose, rather than based on a latent variable, as with the LCCA. This research revealed that visitor groups show different preferences based on a latent variable. The existence of these groups may also explain some of the travel choice patterns that were unclear in previous research. The existence of these groups must therefore be taken into account when conducting research into travel behaviour.

In addition, previous research often did not clearly describe the precise information that was disseminated and how it was communicated (content, medium, time, visualisation, etc.). The message did not seem to be an important part of the analysed studies. Therefore, the influence of these aspects cannot properly be determined, which suggests that communication is not tailored to the travellers' preferences. By implementing the preferences that emerged from this study, some message-related factors that may influence follow-up behaviour can be excluded. By reducing the number of variables, the behavioural patterns and how they can be influenced could become more clear.

In this study, only a few people indicated that they do not want to receive information on their phones at all. The fact that many people are open to receiving travel and event information indicates an opportunity to influence people's choices. Therefore, information provision via mobile phones seems to be a good way to inform people and influence their choices, which shows opportunities for new research.

The information communication preferences of three clusters were presented in this study. The results of this research can already be applied in practice by adapting the communication to these preferences, depending on the target group. When an event organiser can collect personal data from visitors, the information provision can best be adapted to their preferences. When this is not possible, the communication can be aligned with the expected target group or the average person's preferences. When the target group is known, but not their contact information, at least the preferences regarding content and timing can be implemented.

Not only the event organisations can apply the results. Sweco can also use the results to better align the information provision used in CROMAS with the preferences of the people whose travel behaviour they want to influence. If the travellers' contact details are known, the results for the content, timing, and communication channels can be applied. The results show which channels are most interesting for CROMAS to focus on.

There are several remarkable results that are certainly important for event organisations and Sweco to take into account. The first aspect concerns the timing of the message. The vast majority of people indicated that they would like to receive a specific type of information before they leave, but far fewer people were interested in this during the trip or at the location. When an organisation needs to inform people at a later moment, it is better to consider other ways of informing them. If the information is known in advance, sending it to the visitors before departure is best.

The second aspect is related to the communication channels. It is remarkable that, even though event organisations communicate a lot of information via social media, these are not the most popular channels through which people like to receive travel and event information. This was a striking and unexpected result because of the popularity of these platforms. This is an important finding for event organisations to take into account. Many people seem to find that social media are not meant to provide important information. They prefer to receive important information by email.

The third aspect is about who should receive the messages. The results showed that hardly anyone had their travel company completely prepare their trip, so the more people are familiar with the information, the better. However, the *non-Amsterdam students* are most interested in information and are most often willing to act on the information compared to the other clusters. It seems that some people just want to have the information, while the *non-Amsterdam students* want to have it and plan to act on it. Because they are the most impressionable cluster, the information provision must at least be in line with their preferences.

5.3. Limitations

The way a survey is distributed has consequences for the sample that is analysed and the extent to which it represents the Dutch population. Because the survey was distributed online, many responses could be collected in a short time, which was a major advantage. However, the distribution mainly took place among a convenience sample. This ensured that the participants were somewhat younger and more educated than the Dutch population. In addition, many participants live in the Randstad. These differences between the sample and the Dutch population could have been prevented by asking people on the street whether they would like to complete the survey or by distributing the survey via another party. If this had been a better representation of the Dutch population, the clusters might have looked slightly different. Since the average mass event visitor is younger than the average Dutch resident, the biggest difference between the average research participant and the average mass event visitor is related to the level of education. Whether adding this group will result in an extra cluster or whether these people will be included in the existing clusters cannot be concluded from this study. Higher-educated people often have better-paid jobs, increasing their chances of owning a car. The education level-related effects on information communication preferences can not be concluded from this research.

When a survey is set up, there are always a few challenges. It is important that people are not coerced in a particular direction while answering a question, and no interpretational differences are possible. Short event descriptions were given to the participants before they had to answer the event-related questions to limit interpretational differences of an event. However, even with a description, every person has their own experiences, which result in thoughts and feelings about the events, in-

fluencing their choices. The interpretation differences became clear in the results of the second part of the survey, behavioural intention. Some participants indicated that they would act on the message about parking information, while they were not planning to travel by car according to their stated preferred travel mode. The clusters did not rank many messages significantly differently. People did not seem to judge the messages with the idea of travelling by a particular mode of transportation or from a particular 'cluster type view'. This misunderstanding could have been prevented by telling the participants, before starting with this second part of the survey, that they had to judge the messages from the perspective that they were travelling by their stated preferred mode. The differences between the clusters are expected to be more significant than the results show now.

The way questions are answered also depends on the included answer options. The 'other' option showed that apps and websites that provide information about public transport timetables are often used for trip preparation. Because this answer was not one of the possible options, there is a chance that people placed it under a different answer option. Not every participant comes up with the options themselves when certain options are not given, which can influence the results. This could also apply to other questions without the 'other' option. A disadvantage of a survey is that it does not allow participants to substantiate their choices.

The included messages in the survey part about behavioural intention differed based on content, but nothing was specified about the timing of the message and the channel through which it was sent. These are also components on which the preferences of the clusters were determined. To paint a complete picture of the extent to which the preferences match their behaviour, these factors should also be included when people are asked if they are inclined to act on the information.

Finally, during the Latent Class Cluster Analysis, the model failed to show a significant result with all indicators. As explained in the results, a selection of the indicators was included in this study. Some excluded indicators relate to the channels the participants use to look up information, and another excluded indicator is about travelling with a disability. Because the channels on which people prefer to receive information were included as covariates and few people indicated interest in information about travelling with a disability, it was expected that the model could still estimate the clusters properly with the included indicators. However, it did have consequences for how the clusters were determined. With a different combination of indicators, the participants would have been divided over the clusters differently, which would have affected the differences in follow-up behaviour between the clusters, which could have resulted in more or less visible differences between the clusters.

5.4. Recommendations

Several recommendations for follow-up research emerged from this study. The first recommendation resulted from the limited differences in behavioural intention between clusters after seeing the information messages. The results did not show many significant differences between the intentions of the clusters. It is expected that there are more significant differences between the clusters than what has been found in this study, caused by the perspective from which people assessed the message. In addition, only one message was presented for each topic, which was meant to give a first impression of the connection between the indicated preferences and behavioural intentions. No firm conclusions can be drawn from one message. To properly investigate the differences between the clusters, different messages on the same topic will have to be shown to the participants to exclude the effect of elements like wording and visual representation. The results of this are important because they show whether it is actually useful to distinguish different groups of visitors when sending information.

The second aspect that needs more attention is the way in which information is presented. This study considered the content, communication channels, and the moment of receipt. However, precise wording and visualisation of the message are expected to be also important for visitors. This might even differ per communication channel. Previous research has already shown that credibility, reliability, and clear representation of a message are important. Still, not much is known about what this exactly means and what the effects are. Furthermore, this research showed that people have preferences considering the messages' content, timing, and communication channels. So, message-related

aspects seem to influence the extent of follow-up behaviour. To achieve the best result, learning more about these aspects is important. This may ensure that even more people see the message and follow the included advice.

Thirdly, it has become clear that people are most interested in receiving information before they leave. However, reaching people on their way or at their location is sometimes important in case of acute changes. People who are less interested in information at these moments are also less likely to check their message notifications. This is an important finding because these are the moments currently focused on in crowd management. Therefore, it is important to investigate the possibilities of reaching people who have already left home or are already at their location. In this way, the crowds can be better controlled if unexpected things happen or the organisation wishes to inform their visitors later.

Bibliography

Abkowitz, M. D. (1981). An analysis of the commuter departure time decision. *Transportation*, 10(3):283–297.

Ajzen, I. (1991). The theory of planned behavior. *Organizational behavior and human decision processes*, 50(2):179–211.

Allant International University (2015). exploratory factor analysis in SPSS example 01.

Arellana, J., Daly, A., Hess, S., de Dios Ortúzar, J., and Rizzi, L. I. (2012). Development of surveys for study of departure time choice: two-stage approach to efficient design. *Transportation Research Record*, 2303(1):9–18.

Bagloee, S. A., Ceder, A., and Bozic, C. (2014). Effectiveness of en route traffic information in developing countries using conventional discrete choice and neural-network models. *Journal of Advanced Transportation*, 48(6):486–506.

Bakk, Z., Oberski, D. L., and Vermunt, J. K. (2014). Relating latent class assignments to external variables: Standard errors for correct inference. *Political analysis*, 22(4).

Basu, N., Haque, M. M., King, M., Kamruzzaman, M., and Oviedo-Trespalacios, O. (2022). A systematic review of the factors associated with pedestrian route choice. *Transport reviews*, 42(5):672–694.

CBS (2022). CBS Statline.

CBS (2023). Bevolkingspiramide.

Chowdhury, M. Z. I. and Turin, T. C. (2020). Variable selection strategies and its importance in clinical prediction modelling. *Family medicine and community health*, 8(1).

Clark, S. L. and Muthén, B. (2009). Relating latent class analysis results to variables not included in the analysis.

Collins, L. M. and Lanza, S. T. (2009). *Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences*, volume 718. John Wiley & Sons.

Columbia University Mailman School Public Health (2014). Exploratory Factor Analysis.

CROMAS consortium (2022). Cromas Consortium.

Daamen, W. (2004). *Modelling passenger flows in public transport facilities*. PhD thesis, Delft University of Technology.

Daud, K. A. M., Khidzir, N. Z., Ismail, A. R., and Abdullah, F. A. (2018). Validity and reliability of instrument to measure social media skills among small and medium entrepreneurs at pengkalan datu river. *International Journal of Development and sustainability*, 7(3):1026–1037.

DiStefano, C. and Kamphaus, R. (2006). Investigating subtypes of child development: A comparison of cluster analysis and latent class cluster analysis in typology creation. *Educational and Psychological Measurement*, 66(5):778–794.

Donald, I. J., Cooper, S. R., and Conchie, S. M. (2014). An extended theory of planned behaviour model of the psychological factors affecting commuters' transport mode use. *Journal of environmental psychology*, 40:39–48.

Festival Fans (2023). Amsterdam Open Air 2023 - Tickets and Line-up - 3 and 4 juni - Gaasperpark.

Field, A. (2009). *Discovering Statistics Using SPSS*. ISM (London, England). SAGE Publications.

Galama, I. (2015). Route choice behaviour at mass events: Stated-versus revealed preferences of pedestrian route choices at sail amsterdam 2015. Master's thesis, TU Delft.

Gan, H. (2015). To switch travel mode or not? impact of smartphone delivered high-quality multimodal information. *IET Intelligent Transport Systems*, 9(4):382–390.

Gao, F. (2014). Time pressure impact on en-route choice behavior under guidance information. In *Applied Mechanics and Materials*, volume 614, pages 539–542. Trans Tech Publ.

Gao, L., Juan, Z., Ni, A., and Jing, P. (2014). The effect of travel information on travelers' choice of travel modes and routes: a case study of the travel between the campuses. *Mathematical Problems in Engineering*, 2014.

Geng, J., Long, R., and Chen, H. (2016). Impact of information intervention on travel mode choice of urban residents with different goal frames: A controlled trial in xuzhou, china. *Transportation Research Part A: Policy and Practice*, 91:134–147.

Gorsuch, R. L. (1983). *Factor Analysis*. Psychology Press.

Götting, M. C. (2021). Age distribution of festival-goers in Europe 2016 | Statista.

Hair, J. F., Anderson, R. E., Tatham, R. L., and Black, W. C. (1995). *Multivariate Data Analysis (4th Ed.): With Readings*. Prentice-Hall, Inc., USA.

Hanson, S. and Schwab, M. (1986). Describing disaggregate flows: individual and household activity patterns. *Geography of Urban Transportation*.

Harris, L. R. and Brown, G. T. (2010). Mixing interview and questionnaire methods: Practical problems in aligning data. *Practical Assessment, Research, and Evaluation*, 15(1):1.

Heerhugowaard (2018). Evenementenbeleid.

Henson, R. K. and Roberts, J. K. (2006). Use of exploratory factor analysis in published research: Common errors and some comment on improved practice. *Educational and Psychological measurement*, 66(3):393–416.

Heyblom, O. (2022). Topdrukte, maar geen treinen naar Arena: wat kunnen we verwachten?

Hino, Y. and Nagatani, T. (2014). Effect of bottleneck on route choice in two-route traffic system with real-time information. *Physica A: Statistical Mechanics and Its Applications*, 395:425–433.

Islam, M. F. (2018). *Impact of ubiquitous real-time information on bus passenger route choice*. PhD thesis, Edinburgh Napier University.

Jou, R.-C. (2001). Modeling the impact of pre-trip information on commuter departure time and route choice. *Transportation Research Part B: Methodological*, 35(10):887–902.

Khattak, A. J., Schofer, J. L., and Koppelman, F. S. (1995). Effect of traffic information on commuters' propensity to change route and departure time. *Journal of Advanced Transportation*, 29(2):193–212.

Knijn, R. (2020). Maas for commuting trips: Potential for success or deception? Master's thesis, Technische Universiteit Delft.

Kohlhardt, R., Honey-Rosés, J., Fernandez Lozada, S., Haider, W., and Stevens, M. (2018). Is this trail too crowded? a choice experiment to evaluate tradeoffs and preferences of park visitors in garibaldi park, british columbia. *Journal of Environmental Planning and Management*, 61(1):1–24.

Kootstra, G. (2004). Exploratory Factor Analysis Theory and Application.

Kroesen, M. (2022). SEN1721 module 3: Latent class cluster models introduction.

Le, Y., Takeuchi, R., and Aoyagi, S. (2022). An analysis of influencing factors on behavioral change in departure time choice for homeward trips after sightseeing. In *2021 6th International Conference on Intelligent Transportation Engineering (ICITE 2021)*, pages 1017–1026. Springer.

Le-Klähn, D.-T., Roosen, J., Gerike, R., and Hall, C. M. (2015). Factors affecting tourists' public transport use and areas visited at destinations. *Tourism Geographies*, 17(5):738–757.

Liu, E.-H., Tian, S., and Chen, Q. (2015). Mechanism research of bus dynamic information impact on the commuter travel behavior. In *Applied Mechanics and Materials*, volume 743, pages 660–666. Trans Tech Publ.

Liu, S. X. and Guan, H. Z. (2011). The impact of traffic information on drivers' day-to-day route choice behavior. In *Applied Mechanics and Materials*, volume 97, pages 925–930. Trans Tech Publ.

Lu, X., Gao, S., Ben-Elia, E., and Pothering, R. (2014). The effect of real-time information on travelers' route-choice behaviors. *Mathematical Population Studies*, 21(4):205–219.

Luo, R., Fan, Y., Yang, X., Zhao, J., and Zheng, S. (2021). The impact of social externality information on fostering sustainable travel mode choice: A behavioral experiment in zhengzhou, china. *Transportation research part A: policy and practice*, 152:127–145.

Mahdavi, F., Abbaspour, R. A., and Naseri, M. (2022). The impact of access to geospatial information and land-use on users' travel behavior in disruption management in road networks of smart city. *Engineering Journal of Geospatial Information Technology*, 10(1):39–67.

Mayo, F. L. and Taboada, E. B. (2020). Ranking factors affecting public transport mode choice of commuters in an urban city of a developing country using analytic hierarchy process: The case of metro cebu, philippines. *Transportation Research Interdisciplinary Perspectives*, 4:100078.

Molin, E., Mokhtarian, P., and Kroesen, M. (2016). Multimodal travel groups and attitudes: A latent class cluster analysis of dutch travelers. *Transportation Research Part A: Policy and Practice*, 83:14–29.

Neill, J. (2013). Exploratory factor analysis.

Nylund-Gibson, K. and Choi, A. Y. (2018). Ten frequently asked questions about latent class analysis. *Translational Issues in Psychological Science*, 4(4):440.

Osborne, J. (2014). *Best Practices in Exploratory Factor Analysis*. CreateSpace Independent Publishing Platform.

Pahwa, A. (2023). The History Of WhatsApp.

Picheta, R. and Bae, G. (2022). Seoul crowd disaster leaves South Korea reeling, as death toll rises to 154.

Qian, M. and Jiang, J. (2020). Covid-19 and social distancing. *Journal of Public Health*, pages 1–3.

Qualtrics (2022). Complete Guide to Factor Analysis (Updated 2023).

Ramaekers, K., Reumers, S., Wets, G., and Cools, M. (2013). Modelling route choice decisions of car travellers using combined gps and diary data. *Networks and Spatial Economics*, 13(3):351–372.

Ramos, G. d. O., Bazzan, A. L., and da Silva, B. C. (2018). Analysing the impact of travel information for minimising the regret of route choice. *Transportation Research Part C: Emerging Technologies*, 88:257–271.

Reinolsmann, N., Alhajyaseen, W., Brijs, T., Pirdavani, A., Ross, V., Hussain, Q., and Brijs, K. (2022). Delay or travel time information? the impact of advanced traveler information systems on drivers' behavior before freeway work zones. *Transportation research part F: traffic psychology and behaviour*, 87:454–476.

Rodrigues, J. M. F. (2017). What is Mass Events | IGI Global.

Ruesink, C. (2019). 'Clubs moeten van fan-gericht denken naar consument-gericht'.

Schreiber, J. B. (2017). Latent class analysis: an example for reporting results. *Research in social and administrative pharmacy*, 13(6):1196–1201.

Schreiber, J. B. (2021). Issues and recommendations for exploratory factor analysis and principal component analysis. *Research in Social and Administrative Pharmacy*, 17(5):1004–1011.

Segadilha, A. B. P. and da Penha Sanches, S. (2014). Identification of factors that influence cyclistś route choice. *Procedia - Social and Behavioral Sciences*, 160:372–380. XI Congreso de Ingenieria del Transporte (CIT 2014).

Shelby, B., Heberlein, T. A., Vaske, J. J., and Alfano, G. (1983). Expectations, preferences, and feeling crowded in recreation activities. *Leisure Sciences*, 6(1):1–14.

Simons, D., Clarys, P., De Bourdeaudhuij, I., de Geus, B., Vandelanotte, C., and Deforche, B. (2014). Why do young adults choose different transport modes? a focus group study. *Transport policy*, 36:151–159.

Skelley, T., Namoun, A., and Mehandjiev, N. (2013). The impact of a mobile information system on changing travel behaviour and improving travel experience. In *International Conference on Mobile Web and Information Systems*, pages 233–247. Springer.

Statistics Solutions (2021). Factor Analysis.

Steed, J. L. and Bhat, C. R. (2000). On modeling departure-time choice for home-based social/recreational and shopping trips. *Transportation Research Record*, 1706(1):152–159.

Straub, D., Boudreau, M.-C., and Gefen, D. (2004). Validation guidelines for is positivist research. *Communications of the Association for Information systems*, 13(1):24.

Suhr, D. D. (2006). Exploratory or confirmatory factor analysis? *roceedings of the 31st Annual SAS? Users Group International Conference*.

Tabachnick, B. G., Fidell, L. S., and Ullman, J. B. (2013). *Using multivariate statistics*, volume 6. pearson Boston, MA.

Thompson, B. (2004). Exploratory and confirmatory factor analysis: Understanding concepts and applications. *Washington, DC*, 10694(000).

Thorhauge, M., Rich, J., and Cherchi, E. (2015). Departure time choice: Modelling individual preferences, intention and constraints. *Technical University of Denmark*.

TopScriptie (2022). Goede enquêtevragen opstellen.

Toro-González, D., Cantillo, V., and Cantillo-García, V. (2020). Factors influencing demand for public transport in colombia. *Research in Transportation Business & Management*, 36:100514.

Tyrinopoulos, Y. and Antoniou, C. (2012). Factors affecting modal choice in urban mobility. european transport research review, 5 (1), 27-39.

Uebersax, J. (2000). Conditional Dependence in Latent Class Models.

Van Berkum, E. and Van der Mede, P. (1993). The impact of traffic information: dynamics in route and departure time choice. Master's thesis, Technische Universiteit Delft.

van Essen, M., Thomas, T., Chorus, C., and van Berkum, E. (2019). The effect of travel time information on day-to-day route choice behaviour: evidence from a real-world experiment. *Transportmetrica B: Transport Dynamics*, 7(1):1719–1742.

van't Veer, R., Annema, J. A., Araghi, Y., de Almeida Correia, G. H., and van Wee, B. (2023). Mobility-as-a-service (maas): A latent class cluster analysis to identify dutch vehicle owners' use intention. *Transportation Research Part A: Policy and Practice*, 169:103608.

Verizon Editorial Team (2023). The History of Apple iPhones.

Vermunt, J. K. and Magidson, J. (2005). Latent gold® choice 4.0 user's manual. *Statistical Innovations Inc., Belmont, MA*.

Volberda, L. (2020). Analyzing citizens' views on new spatial-infrastructure projects: From the average view towards various clusters within the participatory value evaluation method. Master's thesis, Technische Universiteit Delft.

Wang, D. and Liu, Y. (2015). Factors influencing public transport use: a study of university commuters' travel and mode choice behaviours. In *7th State of Australian Cities Conference, 9-11 December 2015, Gold Coast, Australia*.

Wang, J. and Rakha, H. (2020). Empirical study of effect of dynamic travel time information on driver route choice behavior. *Sensors*, 20(11):3257.

Wang, J. and Rakha, H. A. (2015). Impact of dynamic route information on day-to-day driver route choice behavior. Technical report, Transportation Research Board.

Wang, S. (2015). The function of individual factors on travel behaviour: comparative studies on perth and shanghai. In *Proceedings of the State of Australian Cities Conference*. State of Australian Cities Conference 2015.

Williams, B., Onsman, A., and Brown, T. (2010). Exploratory factor analysis: A five-step guide for novices. *Australasian journal of paramedicine*, 8:1–13.

Xu, C., Wang, W., Yang, C., and Li, Z. (2010). Exploring the impact of real-time travel information on drivers' route choice using stated preference data. In *ICCTP 2010: Integrated Transportation Systems: Green, Intelligent, Reliable*, pages 2214–2221. American Society of Civil Engineers.

Xue, Y., Fan, H., and Guan, H. (2019). Commuter departure time choice considering parking space shortage and commuter's bounded rationality. *Journal of Advanced Transportation*, 2019.

Yan, H., Tao, C., Cui, Y., and Li, Y. (2020). The impact of real-time bus information on passengers' travel choice. In *CICTP 2020*, pages 3024–3037. American Society of Civil Engineers.

A

Scientific paper

Preferences of visitors of mass events towards travel information messages

B.D. den Hollander

Delft University of Technology
April 2023

Abstract. During the influx to mass events in the past, situations have regularly arisen that people experienced as unpleasant or even unsafe. Nowadays, many researchers focus on regulating this influx by influencing travel choices, where only little attention is paid to travellers' preferences. This research identifies the preferences of different type of visitors towards receiving information, and the extent to which visitors with different information profiles act on the information messages they receive. A survey was distributed (N=378), and the number of variables was reduced by performing an Exploratory Factor Analysis. The Latent Class Cluster Analysis distinguished three clusters based on their information message preferences. The results show that people are most interested in receiving information via email a few days before departure and that it is challenging to inform them at a later time. The most impressionable cluster mainly consists of young people, unfamiliar with the event area. Event organisations can apply the results by better-aligning communication with the preferences of their visitors. The differences in behavioural intentions between the clusters turned out to be limited. This must be further investigated since it is not in line with expectations.

Keywords: Travel behaviour, information provision, message preferences, Latent Class Cluster Analysis, Exploratory Factor Analysis, behavioural intention

Introduction

Crowd formation takes place all over the world. Usually, this does not cause any problems, but sometimes situations arise where people feel uncomfortable or even endangered. A reason why so many people come together can be the influx to a mass event. A mass event is a planned gathering of a large group of people for a common purpose at a specific location [1]. If unsafe situations arise, this is often during the influx because many people want to arrive simultaneously. It is essential to spread the visitors more evenly over time and space to prevent this from happening, based on the current (or expected) level of crowdedness. To achieve this, visitors must be informed with travel and event information.

Nowadays, almost everyone owns a mobile phone, which has created new possibilities for informing people. Together with its partners, Sweco developed a tool called Crowd Management System (CROMAS). CROMAS measures the current crowdedness and predicts the prospective crowdedness based on historical and real-time data, visualised in an interactive dashboard. In case the situation does not correspond to the desired situation, measures can be taken by sending people informational messages on their mobile phones or by sending instructions to the traffic controllers [2]. This is an attempt to make people reconsider their travel choices.

In previous research, little has been discovered about the best way to approach people, especially visitors of mass events. To improve the effectiveness of CROMAS, CROMAS must be better attuned to the information preferences of different travellers. This research aimed to identify the preferences of different types of visitors towards receiving information, and to what extent visitors with different information profiles act on the information messages they receive. The focus is on the influx to mass events in Amsterdam, for which the preferences regarding the content, communication channels, and timing of the message received via mobile phones, were examined.

Visitor profiles were identified by performing a Latent Class Cluster Analysis (LCCA), in which visitors are probabilistically assigned to a cluster based on an unobserved variable [3]. After which was examined whether their behavioural intentions aligned with their stated preferences.

The remainder of this paper is structured as follows. First, the methods used, consisting of data collection and data analysis, are discussed in more detail. The data was collected by distributing a survey, the design of which is explained. This is followed by the data analysis, consisting of Exploratory

Factor Analysis (EFA), LCCA, and behavioural intention analysis. The results, including sample data, are discussed in the Results section. The final section provides this research's conclusion, implications, limitations, and recommendations.

Methodologies

Survey design

The research question consists of two parts. First, people's preferences regarding receiving information messages were examined, and after that, it was studied to what extent their preferences corresponded to their behavioural intentions. This distinction is also reflected in the survey. The first part examined the information communication preferences for three different trip purposes (leisure-related mass events in Amsterdam: Music festival (M), Light festival (L), and Football match (F)). In the second part, the behavioural intention was investigated based on five messages in which the participants indicated how likely they would act on these messages. The survey ended with part three, where some personal factors were asked. The survey was designed using Qualtrics and distributed online in Dutch.

The theory has shown that people make their modality, route, and departure time choices based on many different choice factors, which can differ per person. Only a small part of this can be influenced in the short term. The most frequently mentioned factors in the analysed studies about influencing travel choices were included in this research. These factors were expected to influence the choices since these factors showed significant effects in previous studies. In addition, factors not (often) found in previous research have been included, mainly related to the message. As a result, the event-related questions included the factors: message content, trip preparation channels, preferred time to receive information, preferred communication channels, the preferred mode of transport, and trip preparation timing. The personal factors questions included the factors: education level, age, gender, familiarity with Amsterdam, and mode of transport options.

Exploratory Factor Analysis

Three factor analyses were conducted to reduce the number of variables by combining related ones, and to gain insight into the underlying theoretical structure of the data [4]. Factor analysis is based on the idea that there are deeper underlying factors that connect the variables. The Exploratory Factor Analysis was most suitable for this research compared to the Confirmatory Factor Analysis (CFA) since no previous study was found that examined the information message preferences of visitors of mass events, so there was no hypothesis to test. In addition, the observed variable (survey question answer) could potentially be a measure of every factor. The EFA was performed in IBM SPSS Statistics (version 28). Principal Components Analysis was used as a factor extraction method for all analyses, and the orthogonal rotation with the Varimax technique was used.

First, it was checked whether the data were suitable for factor analysis. The data are suitable if the sample size is at least 300 [5], the Kaiser-Meyer-Olkin (KMO) is ≥ 0.50 (measures sample sufficiency for each variable in the model and for the complete model), and the Barlett's Test of Sphericity is significant ($p < 0.05$) (compares the correlations between the observed correlation matrix and the identity matrix) [6]. Furthermore, all the extractions of communalities have to be larger than 0.3.

The number of factors was determined based on the Kaiser rule (eigenvalues > 1) [7], which was verified by the scree plot and parallel analysis (actual values > parallel analysis values) [8].

Finally, the reliability was tested by calculating Cronbach's Alpha values (> 0.6) [9].

Latent Class Cluster Analysis

The LCCA was performed to find groups with similar message preferences. The LCCA was executed in LatentGOLD 5.0 after the number of variables was reduced by performing EFA. The LCCA statistically identifies latent classes within a population where people in the same class have similar preferences [3]. "The goal is to maximise homogeneity within clusters and heterogeneity between clusters" [10].

First, the measurement model was estimated, including only indicators which define the latent variable. In this research, the indicators were the preferred moment of reception, trip preparation moment, trip preparation channels, and preferred message content. These (message-related) factors define the latent variable, which is the attitude towards information messages. The latent variable considers the association between the indicators, which are assumed to be independent.

The measurement model was used to determine the optimal number of classes. To do so, three criteria were checked [11]. The value for both Bayesian Information Criteria (BIC) and Akaike Information Criterion (AIC) needs to be as low as possible since this indicates the most parsimonious model [12]. The third criterion concerns the Bivariate Residuals (BVR), for which a maximum value of 3.84 is allowed. Because a value above 3.84 indicates statistically sig-

nificant indicators at a 5% level. This is undesirable as it would imply a significant relationship between the indicators. If the requirements were not met, the number of indicators had to be reduced until a suitable model was found.

Second, the structural model was estimated, in which the covariates were added to the model, which describe the characteristics of the individuals in that class, also called observed variables. In this study, the covariates were the personal factors that influence the message preferences. These factors do not define the message preferences but might influence them, which is exactly what a covariate does. To start with, all covariates were added to the model. Backward elimination (elimination of the covariate with the smallest effect on the model fit) was applied until all covariates were significant. Significant covariates are indicated by Wald values higher than 3.84 and p-values below 0.05. This is desired since only a significant relation means that the covariate is likely to predict the class membership in the population [10]. The eliminated covariates were made inactive. Inactive covariates do not affect class membership but can be used to describe the classes.

Finally, the model's accuracy was investigated by checking whether the entropy R-squared value was higher than 0.8. Indicating that the LCCA most likely performed best in recovering the true value used in the simulation study, and demonstrated relatively good coverage and power in the examined settings [13].

Behavioural intention

This part of the study investigated whether the indicated preferences of the clusters aligned with their behavioural intention. The results of the second part of the survey were used for this. First, the participants were assigned to a cluster based on the highest probability, leading to "hard partitioning" [14]. After this, the distributions of the scores among the clusters were compared. This gave a first indication of the actual follow-up behaviour. Since only five messages were shown to the participants, no firm conclusions could be drawn, but striking observations may lead to future research.

Results

Descriptive personal results

The survey resulted in a sample size of 378 (response rate 70%). The participants represent the Dutch population quite well based on gender but not so well based on age and level of education [15]. The participants are significantly younger and higher educated compared to the Dutch population. Since the average event visitor is younger than the average Dutch inhabitant, the actual differences are smaller, so the biggest difference between the participants and the Dutch population is based on education level.

Descriptive data results

The general results show a lot of overlap between the different events. The preparation moment distribution is compared with the preferred communication moment distribution in Figure 1. People would like to receive information for all three events before preparing their trip, preferably a few days before the event takes place (a maximum of one week in advance). This can be seen in Figure 1 because the blocks on the left are relatively larger than the blocks on the right side for COM compared with PREP per event. Everyone must receive this information, as hardly anyone leaves the entire preparation to their travel company.

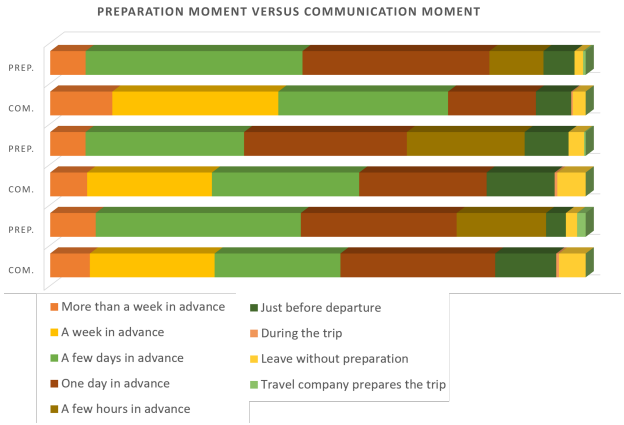

Figure 1: *Preparation moment compared with preferred communication moment per event type (PREP. = preparation moment and COM. = communication moment)*

It is important to provide visitors only with the necessary information and avoid bombarding them with additional information to increase the likelihood of them taking action on the information provided. The desired content is illustrated in Figure 2, with participants being allowed to select multiple answers. Each participant selected an average of three types of information, with recent traffic disruptions, route advice, and event schedule changes as the most popular content.

Because visitors tend to look up less information en route or at the location, as much information as possible (including information about the return trip) must be communicated before departure. This can be seen from the results in Table 1 by looking at "total sources consulted", which are much higher before

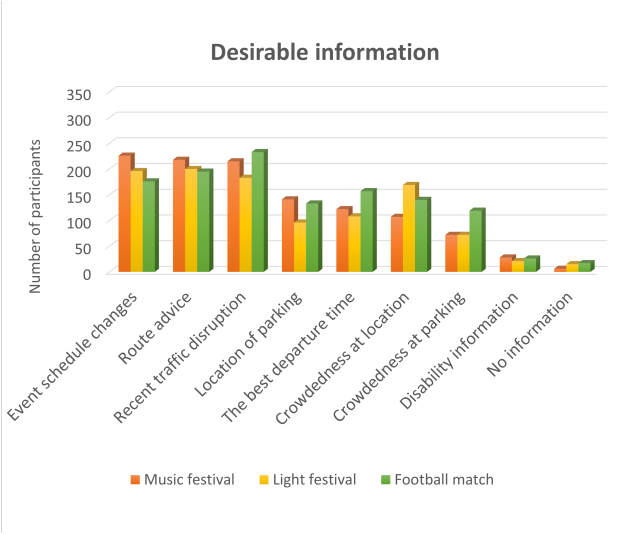

Figure 2: *Desirable travel or event information per event type*

departure than en route, and at the location. These findings are important to take into account if visitors are to be informed. Since this data show many more possibilities for influencing before departure, the rest of this study focuses on before-departure data. Before departure, the organisation's website, emails, and navigation systems are the most consulted channels, and social media are not very popular.

Table 1: *Preparation channels per event type and moment, where M = Music festival, L = Light festival, F= Football match*

	Before departure			En route			At location		
	M	L	F	M	L	F	M	L	F
The organisation's website	303	309	236	46	48	35	49	60	26
Event organiser emails	268	196	218	38	22	36	31	15	19
SM channels of the organisation	127	118	98	47	47	41	43	40	34
SM channels of other organisations	44	47	39	26	21	23	19	19	15
organisation's app (if available)	95	97	91	64	51	59	131	108	77
Navigation systems such as Google Maps	207	193	225	257	231	227	66	101	68
Total sources consulted	1044	960	907	478	420	421	339	343	239

To inform people, it is important to know which channels people use to prepare their trip and through which channels they prefer to receive information messages. The results are visualised in Figure 3, in which participants were allowed to give multiple answers. The most notable value is for email, which is considerably higher than the other values. Other high values can be seen for WhatsApp, event apps, and SMS. Even though event organisations increasingly focus on social media, these are not the channels through which people like to receive information.

Exploratory Factor Analysis

The EFA was performed trice, in which no variables had to be deleted to find a suitable result. The criteria were checked, with the result that the 6 variables

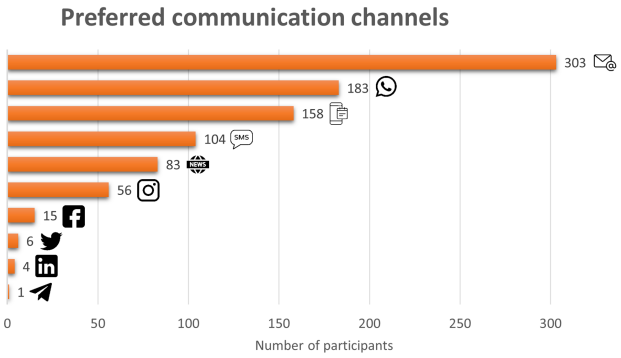

Figure 3: *Preferred communication channels*

related to the moment of communication were combined in one factor, the 24 variables related to content were combined in 7 factors, and the 18 variables related to communication channels were combined in 4 factors. The variables for different events but with the same topic were mostly aggregated. In addition, some topics were merged, such as the organisation's and third parties' social media channels. The factors were used as indicators in the LCCA.

Latent Class Cluster Analysis

A suitable model was found with seven indicators, including six content-related indicators (parking, traffic disruption, route advice, departure time advice, event schedule, and level of crowdedness) and one indicator about timing. Backward elimination started with 22 active covariates and resulted in a significant model, including 13 active covariates (age, gender, familiarity with Amsterdam, 4 travel mode options (own car, shared car, public transport, bicycle), 4 preferred communication channels (WhatsApp, Instagram, news, SMS), and preferred travel mode to music and light festival), and 9 inactive covariates (level of education, 1 travel mode option (P+R), 6 preferred communication channels (Facebook, LinkedIn, Twitter, Telegram, Event Apps, Email), and preferred travel mode to football match).

The LCCA resulted in three clusters, each with its own information preferences and personal characteristics. The results can be seen in Table 2. Furthermore, the information provision preferences of the average person and the three clusters are visualised in Figure 4.

The average person is 37 years old, not more often male or female, highly educated (52% HBO, WO master or higher), visits Amsterdam a few times a year or more often, and has the option to travel to Amsterdam by public transport (94%), by their own car (61%), or via a P+R location (51%). The average person prefers to receive messages about recent

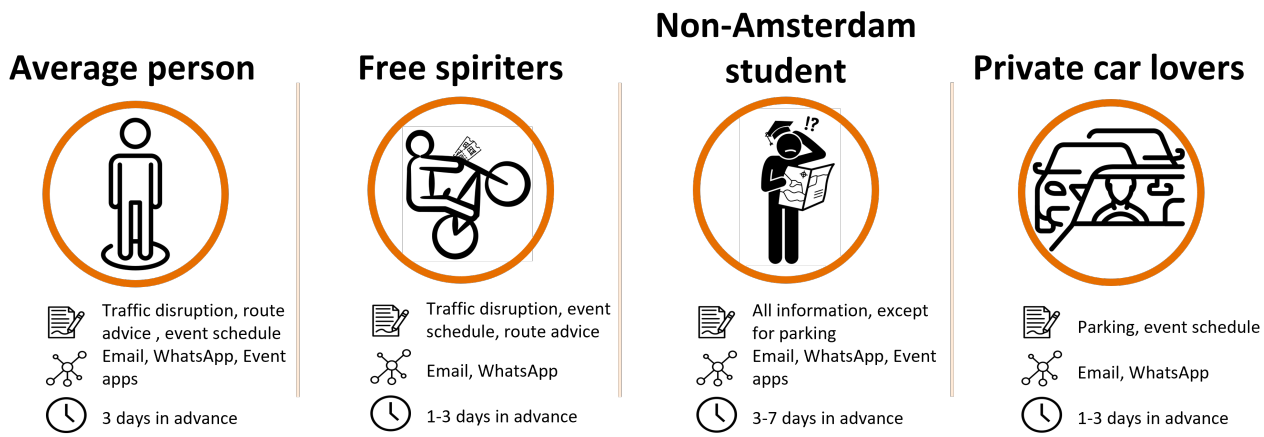

Figure 4: *Information provision preferences per cluster*

traffic disruptions (indicated by 56%), route advice (indicated by 54%), and/or event schedule changes (indicated by 53%). Ideally, the average person receives these messages via email (indicated by 80%) and otherwise through WhatsApp messages (indicated by 48%) before trip preparation. For 93% of the participants, this means one day in advance or more, but with a maximum of one week in advance.

The extent to which each cluster is interested in receiving certain information is shown in Figure 5. The values indicate how often someone indicated to be interested in that type of information. The participants could answer this question with yes or no for all three events. The meaning of the numbers is as follows:

#1	Parking information
#2	Traffic disruption
#3	Route advice
#4	Departure time
#5	Event schedule changes
#6	Crowdedness

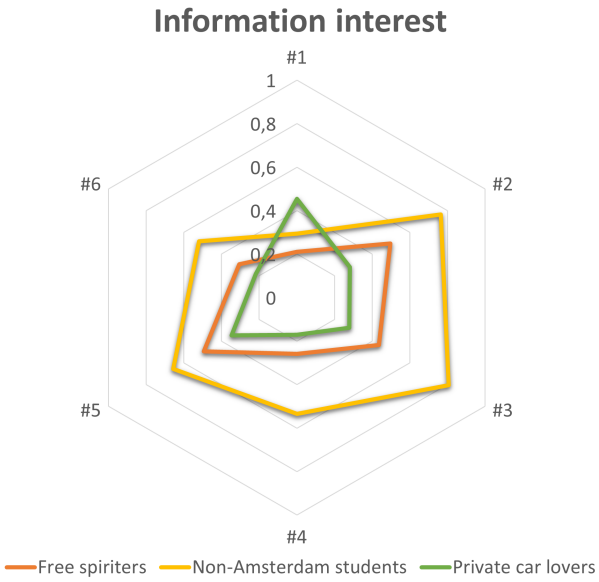

Figure 5: *Interest in certain information content per cluster, frequency on the range 0 (never) to 1 (always)*

The free spiriters (49% of the participants) have a less pronounced preference for a mode of transport than the other clusters, but prefer travelling by public transport (64%) or bicycle (20%), and mostly not the car (7%). They are relatively familiar with Amsterdam, mainly men (62%), and they are highly educated (56% HBO, WO master or higher). They would like to receive information messages about recent traffic disruptions, event schedule changes, and/or route advice by email between 1 and 3 days in advance.

The non-Amsterdam students (35% of the participants) are unfamiliar with Amsterdam, and they are still very young (almost half of the group is between 15 and 25 years old). Therefore, they often have not finished school yet, meaning they are less educated. They crave information and are not very critical of the content and communication channels. However, they prefer to receive messages containing route advice by email between 3 and 7 days in advance. This cluster is the most impressionable cluster.

The private car lovers (16% of the participants) are mainly men (61%) who regularly visit Amsterdam, and they are creatures of habit. They almost all own a car (95%) and really prefer to use it to drive to the events compared with the other clusters (64% compared with 7% and 11%). Because they have their habits, they are not very interested in information, except parking information. They prefer to receive this information by email between 1 and 3 days in advance. This cluster is the least impressionable cluster.

Table 2: *Cluster profiles, including indicators, active covariates, and inactive covariates*

Cluster name		**Free spiriters**	**Non-Amsterdam students**	**Private car lovers**
Cluster size		49%	35%	16%
INDICATORS (MEAN)				
	#1 Parking information	0,2099	0,2949	0,4548
	#2 Traffic disruption	0,4972	0,7645	0,2814
	#3 Route advice	0,4356	0,8062	0,2785
	#4 Departure time	0,259	0,5338	0,1703
	#5 Event schedule changes	0,494	0,656	0,3473
	#6 Crowdedness	0,306	0,5194	0,218
	Moment	3,3283	2,6747	3,5211
ACTIVE COVARIATES				
Age	15-25	29,2%	47%	23%
	26-35	22,2%	24%	25%
	36-45	8,1%	2%	15%
	46-55	17,3%	16%	23%
	56-65	19,5%	10%	13%
	66+	2,7%	2%	2%
	Mean	39	33	39
Gender	Male	62%	34%	61%
	Female	38%	65%	36%
	Other	0%	0%	3%
	Blank	0%	1%	0%
Familiarity with Amsterdam	I live there	10%	0%	3%
	I lived there	11%	2%	10%
	Visit once a month	18%	7%	18%
	Visit few times a year	49%	59%	44%
	Been few times	12%	32%	23%
	Never been	1%	0%	2%
Possible to go to Amsterdam by	Own car	57%	53%	95%
	Shared car	12%	33%	30%
	Public Transport	95%	99%	80%
	Bicycle	20%	5%	23%
Prefer receiving information via	WhatsApp	44%	58%	42%
	Instagram	5%	28%	17%
	News	14%	32%	25%
	SMS	26%	37%	9%
Preferred transport mode M	Own car	4%	6%	57%
	Shared car	0%	1%	3%
	PT	64%	85%	35%
	P+R	10%	1%	2%
	Bicycle	23%	8%	3%
Preferred transport mode L	Own car	2%	7%	75%
	Shared car	1%	2%	0%
	PT	67%	83%	19%
	P+R	10%	1%	2%
	Bicycle	21%	8%	5%
INACTIVE COVARIATES				
Education	Primary education	0%	1%	0%
	Secondary education	9%	17%	2%
	MBO, HBO, WO bachelor	35%	39%	38%
	HBO, WO master	51%	42%	57%
	PhD or higher	5%	2%	3%
	Not educated/unknown	0%	1%	0%
Possible to go to Amsterdam by	P+R	44%	55%	65%
Prefer receiving information via	Facebook	2%	5%	10%
	LinkedIn	1%	0%	5%
	Twitter	2%	1%	3%
	Telegram	1%	0%	0%
	Event Apps	37%	52%	33%
	Email	80%	86%	68%
Preferred transport mode F	Own car	15%	20%	59%
	Shared car	1%	6%	2%
	PT	61%	63%	30%
	P+R	7%	1%	0%
	Bicycle	16%	9%	10%

Behavioural intention

The participants indicated how likely they will act on this information for five information messages. The results of this can be seen per cluster in Figure 6. The results show an unexpected amount of overlap between the different clusters. Although the clusters showed different information preferences, the differences between these results remain minimal. It seems that people did not judge the messages with the idea of travelling by a particular mode of transportation or from a specific 'cluster type view'. However, a significant difference can be seen in the message about the public transport timetable between the *private car lovers* compared with the *free spiriters*, and the *non-Amsterdam students*.

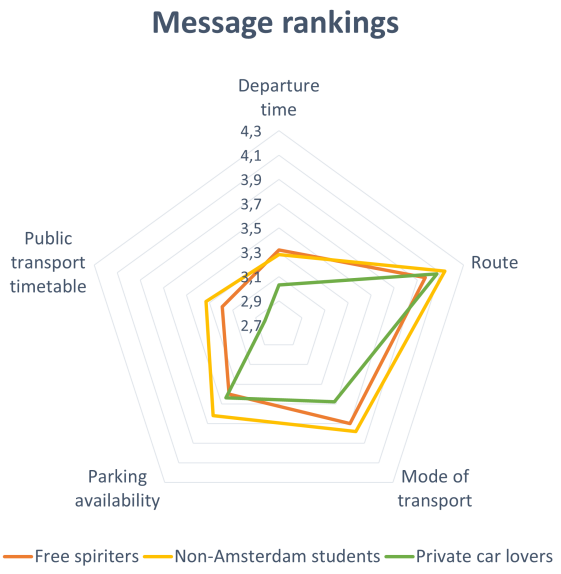

Figure 6: *Message rankings per cluster, scale ranges from 1 (absolutely not) to 5 (definitely)*

No firm conclusions can be drawn from these results because only five messages with different topics were shown to the participants. To draw firm conclusions, multiple messages on the same topic must be presented to the participants to see the influence of other aspects. The results of this research provide a first indication.

Conclusions and recommendations

Conclusion

In recent years, more research has been done into influencing travel behaviour. However, little attention was paid to the needs of the travellers. This research has contributed to this by identifying the preferences of different types of visitors towards receiving information and the extent to which visitors with different information profiles act on the information messages they receive.

The general results clearly showed that sharing as much information as possible with people before departure is essential. In addition, event organisations are increasingly focusing on social media, but the participants expressed a clear preference for communication via email. The results also showed that the information communication preferences for the three included events differed only a little.

The LCCA distinguished three clusters based on preferences towards receiving information. The non-Amsterdam students turned out to be the most impressionable cluster. In addition, they are very interested in receiving a lot of information, which directly offers the opportunity to influence their choices. The private car lovers are least impressionable, except when it comes to car-related information. The free spiriters are familiar with Amsterdam, which makes them particularly interested in information about changes. The behavioural intention in this study did not correspond much with the differences in information preferences of the clusters, which showed that further research is needed to find out what caused these results.

Implications

In this study, clusters emerged that were not distinguished in this way before. The existence of these clusters and the apparent effect of different aspects of the messages on information preferences appear to be important elements to keep in mind in future research on travel behaviour.

The results of this research are relatively easy to implement in practice. Event organisations and Sweco can better tailor communication with visitors to the target group's preferences. This does not require large investments, which makes short-term implementation possible. The most important aspects are that as much information as possible is to be sent via email a few days before the event and not via social media, where it is most important that the non-Amsterdam students receive the information that interests them.

Limitations

The survey was distributed online among a convenience sample. As a result, the sample is higher educated than the average Dutch population. It has not become clear from this study whether a new cluster will emerge if the population of the Netherlands is better represented or whether these people will spread over the existing clusters. Second, there

seem to have been interpretational differences when assessing the messages in the survey. This misunderstanding could have been prevented by telling the participants, before starting with this second part of the survey, that they had to judge the messages from the perspective that they were travelling by their stated preferred mode. In addition, the connection between the information preferences and the behavioural intention would have been more complete if the timing and communication channel would also have been specified, instead of the focus being on the content only.

Recommendations

Several recommendations for follow-up research emerged from this study. The first recommendation for follow-up research resulted from the limited differences in behavioural intention between clusters after seeing the information messages. It is expected that there are more significant differences between the clusters than what was found in this study, caused by the perspective from which people assessed the message. It is important to investigate this more extensively to understand the importance of information preferences.

The second aspect that needs more attention is the way information is presented. This was not often included in previous studies. Still, based on the results of this research, it is expected that more message-related factors than content, timing, and communication channels affect travel choices. Learning more about these aspects is essential to get the best effect.

Thirdly, the results showed that people are most interested in receiving information before departure. However, if important information still needs to be communicated later, it must be investigated how people can best be reached when they are en route or already at the location. It is still hard to inform people at later moments since they look up less information and are therefore less easily accessible.

References

[1] J. M. F. Rodrigues. *What is Mass Events | IGI Global*. Nov. 2017. URL: https://www.igi-global.com/dictionary/mass-events/61900.

[2] CROMAS consortium. *Cromas Consortium*. 2022. URL: https://cromas.io/.

[3] Linda M. Collins and Stephanie T. Lanza. *Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences*. Vol. 718. John Wiley & Sons, 2009.

[4] Statistics Solutions. *Factor Analysis*. Aug. 2021. URL: https://www.statisticssolutions.com/free-resources/directory-of-statistical-analyses/factor-analysis/.

[5] Barbara G Tabachnick, Linda S Fidell, and Jodie B Ullman. *Using multivariate statistics*. Vol. 6. pearson Boston, MA, 2013.

[6] Brett Williams, Andrys Onsman, and Ted Brown. "Exploratory factor analysis: A five-step guide for novices". In: *Australasian journal of paramedicine* 8 (2010), pp. 1–13.

[7] A. Field. *Discovering Statistics Using SPSS*. ISM (London, England). SAGE Publications, 2009. ISBN: 9781847879073. URL: https://books.google.nl/books?id=a6FLF1YOqtsC.

[8] Bruce Thompson. "Exploratory and confirmatory factor analysis: Understanding concepts and applications". In: *Washington, DC* 10694.000 (2004).

[9] Khairul Azhar Mat Daud et al. "Validity and reliability of instrument to measure social media skills among small and medium entrepreneurs at Pengkalan Datu River". In: *International Journal of Development and sustainability* 7.3 (2018), pp. 1026–1037.

[10] Maarten Kroesen. *SEN1721 module 3: Latent class cluster models introduction*. 2022.

[11] Eric Molin, Patricia Mokhtarian, and Maarten Kroesen. "Multimodal travel groups and attitudes: A latent class cluster analysis of Dutch travelers". In: *Transportation Research Part A: Policy and Practice* 83 (2016), pp. 14–29.

[12] Jeroen K Vermunt and Jay Magidson. "Latent GOLD® choice 4.0 user's manual". In: *Statistical Innovations Inc., Belmont, MA* (2005).

[13] Christine DiStefano and RW Kamphaus. "Investigating subtypes of child development: A comparison of cluster analysis and latent class cluster analysis in typology creation". In: *Educational and Psychological Measurement* 66.5 (2006), pp. 778–794.

[14] Zsuzsa Bakk, Daniel L Oberski, and Jeroen K Vermunt. "Relating Latent Class Assignments to External Variables: Standard Errors for Correct Inference." In: *Political analysis* 22.4 (2014).

[15] CBS. *Bevolkingspiramide*. 2023. URL: https://www.cbs.nl/nl-nl/visualisaties/dashboard-bevolking/bevolkingspiramide.

B

Literature overview: Influencing travel behaviour

Table B.1: Overview of literature related to influencing travel behaviour by using information (part 1)

Goal	Area of influence	Message type	Results	Source
"Determining the impact of information intervention on travel mode choice of urban residents with different goal frames"	Mode choice	Information intervention by i.e. television and newspapers	"Consistent and distinct impacts on travel mode choice by clusters, embodied in the simultaneous and significant increase in travel times by green modes. The distinctness of the impacts is that information have a more effective influence on subjects with gain goal frames because their travel times by all three green modes greatly improved."	(Geng et al., 2016)
"Investigating commuters' en-trip mode decision about switching from 'auto' to 'park-and-ride' (P + R)"	Mode choice	High-quality Smartphone delivered multimodal information (SMMI)	"SMMI can significantly influence mode choice and its impacts depend on traveller attributes, driver's previous experience, and level of service attributes. Statistically significant explanatory variables in the model are delay for auto, comfort level of rail transit, gender, education level, income, driving experience, driving frequency, main criterion of mode choice, owning an easy public transportation ride card, previous use of P + R, perceived value of existing real-time traveller information and frequency of using real-time traveller information."	(Gan, 2015)
"Using social externality information to foster sustainable travel mode choice of commuters"	Mode choice	Information interventions about the higher emissions impact of cars	"While these interventions are initially highly effective in reducing intended car usage, the effect becomes less as respondents learn of the level of local air pollution and the risks of outdoor pollution exposure."	(Luo et al., 2021)
"Investigating the impact which mobile information systems can have on travelling behaviour of commuters"	Mode choice	Mobile application which provides bus timetable information	"The results show that transport information did influence the behaviour of travellers, who started trusting the system and used it to target specific busses and thus reduce their bus waiting times."	(Skelley et al., 2013)
"Investigating the impact of real-time bus information on passengers' travel choice"	Mode choice	Public transport real-time information	"The factors that significantly affected the travel choice behavior of bus passengers include (from strong to weak) waiting time, the situation that the family owns the electric bicycle, the degree of congestion in the bus, the frequency of using real-time information, age, occupation."	(Yan et al., 2020)
"Gather the information of the factors that affect travel choices and the extent of such effects both in general cases and when prior information is given"	Mode and route choice	Travel information	"For sanguine travelers, the ranking of the factors is fare > comfort > stability > time and convenience. For choleric and phlegmatic travelers, the ranking is stability > convenience > time > comfort > fare. For all the travelers, the ranking is stability > convenience > fare, time, and comfort."	(Gao et al., 2014)
"Analysing the impact of travel information for minimising the regret of route choice"	Route choice	information provided by a mobile navigation app	"Drivers aim to minimise travel costs. Experiments showing regret is minimized and system converges to the User Equilibrium."	(Ramos et al., 2018)
"Generating insights into the effect of travel information on day-to-day route choice behaviour"	Route choice	Travel time information	"The provision leads to a decline in switching propensity and a higher probability that the shortest route is chosen. Travel time information seems to influence travellers' propensity to shift from one profile to another across different OD-pairs."	(van Essen et al., 2019)
"Let drivers switch routes in case of a bottleneck by real-time information"	Route choice	Real-time information about a bottlenck	"The traffic behavior changes greatly by the competition between the bottleneck's jam and driver's avoidance to the jam. The dynamic transition occurs between the oscillating jam and the stationary jam."	(Hino and Nagatani, 2014)
"Study the effect of travel time information on day-to-day driver route choice behavior"	Route choice	Dynamically updated travel time information	"Historical travel time information enhances behavioral rationality by 10% on average and reduces inertial tendencies to increase risk seeking in the gain domain. Expected travel time information is demonstrated to be more effective than travel time variability information in enhancing rational behavior when drivers have limited experiences."	(Wang and Rakha, 2020)

Table B.2: Overview of literature related to influencing travel behaviour by using information (part 2)

Goal	Area of influence	Message type	Results	Source
"To evaluate the effect of travel time information on route choice behavior"	Route choice	The average and variance travel times from historical data	"The faster route is demonstrated to be more attractive under high variability conditions when drivers lack experience. The effect of information differs by participant and O-D trip at an individual level, which is, to a large extent, dependent on personal traits and route characteristics."	(Wang and Rakha, 2015)
"Learning about the effect of real-time information on Travelers' route choice behaviors"	Route choice	Real-time information regarding a probable incident	"Real-time information significantly reduces network travel time and its variability. However, no equilibrium was observed in any of the experiment sessions."	(Lu et al., 2014)
"Investigating the influence of different traffic information on drivers' day-to-day route choice behavior based on microscopic simulation"	Route choice	3 options: no information, historical information and predictive information	"The impacts of traffic information on drivers are related to the random level of driver's route choice and reliance on the information. In addition, the road network cannot reach user equilibrium in three kinds of information."	(Liu and Guan, 2011)
"Investigating the impact of Bus Dynamic Information on the Commuter Travel Behavior"	Route choice	Dynamic bus information	"Public transport information is closely related to the travel routes choice and travel activities, commuters after obtaining, travel routes will make adjustments."	(Liu et al., 2015)
"Determine the potential interplay among real-time travel time information, socio-demographics of drivers and route choice."	Route choice	Real-time travel information (travel time, travel time variation and the extent of road congestion)	"It can be discovered that (1) real-time travel information has significant effects on drivers' route choice behavior. (2) Drivers pay more attention to travel time information than travel time variation. (3) Among socio-demographics information of drivers, drivers' age, driving experience and gender have considerable correlations with drivers' route choice behavior."	(Xu et al., 2010)
"Investigating the impact of advanced traveler information systems on drivers' behavior before freeway work zones"	Route choice	Advanced Traveler Information Systems (ATIS) (real-time information about the delays and travel times en-route and an alternative route)	"The results showed that the display of zero delays for a detour did influence 74–83% of the drivers to take the alternative route when being displayed on a VMS and a GRIP with free flow attribute framing. When displaying equal total travel times, the GRIP did influence 25% more drivers to follow the alternative route than the VMS."	(Reinolsmann et al., 2022)
"Investigating the impact of ubiquitous real-time information on bus passenger route choice"	Travel choices	Ubiquitous real-time passenger information (URTPI)	"The study reveals that trip length, passenger age and profession are the main factors influencing the use of URTPI. Having access to URTPI, the frequency of its use is strongly influenced by the attributes of information and social norms. Changing time of departure from the start and the boarding time are the two most popular actions."	(Islam, 2018)
"Investigating the impact of access to geospatial information and land-use on users' travel behavior in disruption management in road networks of smart city"	Travel choices	Three scenarios "no information", "advanced information" or "advanced information considering specific landuses"	"The results show that having access to information can lead to a reduction in total travel time, a reduction in potential passenger delays, and an increase in passenger satisfaction in a network disruption."	(Mahdavi et al., 2022)
"Investigating the impact of time pressure on en-route choice behavior under guidance information"	Travel choices	Guidance information via advanced traffic information system (ATIS)	"Under tight time pressure constraints, the degree of influence of disseminated information in route choice behavior is less pronounced. Longer deliberation time frames allow for more elaborate deliberation processes."	(Gao, 2014)
"Investigating the effectiveness of en route traffic information in developing countries using conventional discrete choice and neural-network models"	Travel choices	Radio Traffic Information	RTI must include advice that explicitly suggests alternative routes to the drivers. The frequency of traffic advice in the morning has to be higher than other information to get people to their destination on time. Young drivers constitute majority of the RTI demand.	(Bagloee et al., 2014)
"Investigate the impact of pre-trip information on auto commuters' departure time and route choice"	Departure time and route choice	Pre-trip information	"The results imply that pre-trip information has a different influence on both departure time and route latent variables. The results suggest that younger commuters tend to switch than older ones for both departure time and route switching decisions. Male commuters are more likely to switch than females for both departure time and route switching decisions."	(Jou, 2001)
"Evaluating the effect of traffic information on commuters' propensity to change route and departure time"	Departure time and route choice	Traffic reports containing real-time traffic information	More than 60% of the respondents had used traffic information to modify their travel decisions. Individuals were more likely to use traffic reports for their route changes if they perceived traffic reports to be accurate and timely, and frequently listened to traffic reports.	(Khattak et al., 1995)
"Develop and validate a methodology to predict or estimate the influence of different, new information systems for car drivers"	Departure time and route choice	Motorized information, a discription of components of the traffic system	"A clear structure in the messages results in more adaptations. Furthermore, information reduces the uncertainty about the outcome of choices."	(Van Berkum and Van der Mede, 1993)

C

Survey

Beste reiziger,

Allereerst wil ik je bedanken voor je deelname aan dit onderzoek. Ik ben een student van de master Transport, Infrastructure and Logistics aan de TU Delft. Dit onderzoek is onderdeel van mijn afstudeerthesis onder leiding van de TU Delft en Sweco.
Het doel van mijn onderzoek is om inzicht te krijgen in de voorkeuren van reizigers voor het ontvangen van reis en evenement informatie en de bereidheid om iets met deze informatie te doen. Dit heeft invloed op de mate van drukte onderweg en op locatie.

Het onderzoek duurt ongeveer 7 minuten en bestaat uit drie onderdelen. Het onderzoek is anoniem, er wordt vertrouwelijk met deze informatie om gegaan.

Met vriendelijke groet,
Babette den Hollander

Deel 1 - Evenement gerelateerde vragen
Hierna volgen er een aantal vragen die gaan over drie verschillende evenementen. Dit betekent dat de vragen drie keer langs komen, voor elk evenement één keer. Elk evenement kent een grote opkomst, wat betekent dat er veel drukte zal zijn bij de toestroom en ter plaatse.
De evenementen waar dit onderzoek op focust zijn:

- Muziekfestival in Amsterdam
- Lichtkunstfestival in de straten van Amsterdam
- Eredivisie voetbalwedstrijd in Amsterdam

Stel je bij het beantwoorden van de volgende vragen voor dat je het volgende evenement bezoekt:

Muziekfestival in Amsterdam, een grootschalig, betaald evenement (meer dan 5.000 bezoekers) waarvoor je van tevoren een kaartje hebt gekocht. De tijd waarop je gaat mag je zelf bepalen.

Q1 Welk vervoersmiddel heeft je voorkeur als je reist naar een **Muziekfestival in Amsterdam**?

○ Auto (eigen bezit)
○ Auto (makkelijk te lenen of abonnement voor shared car service)
○ Openbaar Vervoer
○ Auto tot P+R locatie, vanaf daar met het Openbaar Vervoer
○ Fiets, e-bike of scooter (eigen bezit of deelmobiliteit)

Q2 Wanneer bereid je je reis voor als je reist naar een **Muziekfestival in Amsterdam**?

○ Langer dan een week van tevoren
○ Een aantal dagen van tevoren
○ Een dag van tevoren
○ Een paar uur van tevoren
○ Net voor vertrek
○ Ik vertrek zonder voorbereiding
○ Mijn reisgezelschap bereidt de reis voor

Q3 Welke bronnen gebruik je om reis en evenement informatie op te zoeken (1) voor je vertrekt, (2) onderweg, (3) op locatie van een **Muziekfestival in Amsterdam**? (meerdere antwoorden mogelijk, lege rijen zijn toegestaan) (rastervraag)

○ De website van de organisatie
○ E-mails van evenement organisatie
○ Social media kanalen van de organisatie
○ Social media kanalen van andere organisaties
○ App van de organisatie (indien beschikbaar)
○ Navigatiesystemen als Google Maps
○ Anders, namelijk ..

Q4 Wat voor informatie zou je graag willen ontvangen voordat je gaat reizen naar een **Muziekfestival in Amsterdam**? (meerdere antwoorden mogelijk)

○ De beste vertrektijd
○ Routeadvies
○ Drukte ter plaatse
○ Drukte parkeergelegenheid
○ Locatie van parkeergelegenheid
○ Recente verkeershinder (file, ongeval, uitval OV, vertraagd OV)
○ Wijzigingen evenementen schema
○ Extra informatie met betrekking tot reizen met beperking
○ Geen informatie
○ Anders, namelijk ..

Q5 Wanneer wil je reis informatie ontvangen over een **Muziekfestival in Amsterdam**?

○ Langer dan 1 week van tevoren
○ Vanaf 1 week van tevoren
○ Vanaf 3 dagen van tevoren
○ Vanaf de dag van tevoren
○ Vlak voor vertrek
○ Tijdens de reis
○ Ik wil geen informatie ontvangen

Stel je bij het beantwoorden van de volgende vragen voor dat je het volgende evenement bezoekt:

Lichtkunstfestival in de straten van Amsterdam, een gratis, grootschalig, meerdaags evenement. De tijd waarop je gaat en de precieze startlocatie mag je zelf bepalen.

Q6 Welk vervoersmiddel heeft je voorkeur als je reist naar een **Lichtkunstfestival in de straten van Amsterdam**?

○ Auto (eigen bezit)
○ Auto (makkelijk te lenen of abonnement voor shared car service)

○ Openbaar Vervoer
○ Auto tot P+R locatie, vanaf daar met het Openbaar Vervoer
○ Fiets, e-bike of scooter (eigen bezit of deelmobiliteit)

Q7 Wanneer bereid je je reis voor als je reist naar een **Lichtkunstfestival in de straten van Amsterdam**?

○ Langer dan een week van tevoren
○ Een aantal dagen van tevoren
○ Een dag van tevoren
○ Een paar uur van tevoren
○ Net voor vertrek
○ Ik vertrek zonder voorbereiding
○ Mijn reisgezelschap bereidt de reis voor

Q8 Welke bronnen gebruik je om reis en evenement informatie op te zoeken (1) voor je vertrekt, (2) onderweg, (3) op locatie van een **Lichtkunstfestival in de straten van Amsterdam**? (meerdere antwoorden mogelijk, lege rijen zijn toegestaan) (rastervraag)

○ De website van de organisatie
○ E-mails van evenement organisatie
○ Social media kanalen van de organisatie
○ Social media kanalen van andere organisaties
○ App van de organisatie (indien beschikbaar)
○ Navigatiesystemen als Google Maps
○ Anders, namelijk ..

Q9 Wat voor informatie zou je graag willen ontvangen voordat je gaat reizen naar een **Lichtkunstfestival in de straten van Amsterdam**? (meerdere antwoorden mogelijk)

○ De beste vertrektijd
○ Routeadvies
○ Drukte ter plaatse
○ Drukte parkeergelegenheid
○ Locatie van parkeergelegenheid
○ Recente verkeershinder (file, ongeval, uitval OV, vertraagd OV)
○ Wijzigingen evenementen schema
○ Extra informatie met betrekking tot reizen met beperking
○ Geen informatie
○ Anders, namelijk ..

Q10 Wanneer wil je reis informatie ontvangen over een **Lichtkunstfestival in de straten van Amsterdam**?

○ Langer dan 1 week van tevoren
○ Vanaf 1 week van tevoren
○ Vanaf 3 dagen van tevoren
○ Vanaf de dag van tevoren
○ Vlak voor vertrek
○ Tijdens de reis
○ Ik wil geen informatie ontvangen

Stel je bij het beantwoorden van de volgende vragen voor dat je het volgende evenement bezoekt:

Eredivisie voetbalwedstrijd in Amsterdam, een grote, belangrijke wedstrijd (betaald) waarvoor je al een kaartje hebt gekocht en absoluut op tijd wilt komen. Er is dus een vaste starttijd.

Q11 Welk vervoersmiddel heeft je voorkeur als je reist naar een **Eredivisie voetbalwedstrijd in Amsterdam**?

○ Auto (eigen bezit)
○ Auto (makkelijk te lenen of abonnement voor shared car service)
○ Openbaar Vervoer
○ Auto tot P+R locatie, vanaf daar met het Openbaar Vervoer
○ Fiets, e-bike of scooter (eigen bezit of deelmobiliteit)

Q12 Wanneer bereid je je reis voor als je reist naar een **Eredivisie voetbalwedstrijd in Amsterdam**?

○ Langer dan een week van tevoren
○ Een aantal dagen van tevoren
○ Een dag van tevoren
○ Een paar uur van tevoren
○ Net voor vertrek
○ Ik vertrek zonder voorbereiding
○ Mijn reisgezelschap bereidt de reis voor

Q13 Welke bronnen gebruik je om reis en evenement informatie op te zoeken (1) voor je vertrekt, (2) onderweg, (3) op locatie van een **Eredivisie voetbalwedstrijd in Amsterdam**? (meerdere antwoorden mogelijk, lege rijen zijn toegestaan) (rastervraag)

○ De website van de organisatie
○ E-mails van evenement organisatie
○ Social media kanalen van de organisatie
○ Social media kanalen van andere organisaties
○ App van de organisatie (indien beschikbaar)
○ Navigatiesystemen als Google Maps
○ Anders, namelijk ..

Q14 Wat voor informatie zou je graag willen ontvangen voordat je gaat reizen naar een **Eredivisie voetbalwedstrijd in Amsterdam**? (meerdere antwoorden mogelijk)

○ De beste vertrektijd
○ Routeadvies
○ Drukte ter plaatse
○ Drukte parkeergelegenheid
○ Locatie van parkeergelegenheid
○ Recente verkeershinder (file, ongeval, uitval OV, vertraagd OV)
○ Wijzigingen evenementen schema
○ Extra informatie met betrekking tot reizen met beperking
○ Geen informatie
○ Anders, namelijk ..

Q15 Wanneer wil je reis informatie ontvangen over een **Eredivisie voetbalwedstrijd in Amsterdam**?

○ Langer dan 1 week van tevoren
○ Vanaf 1 week van tevoren
○ Vanaf 3 dagen van tevoren
○ Vanaf de dag van tevoren

○ Vlak voor vertrek
○ Tijdens de reis
○ Ik wil geen informatie ontvangen

Q16 Via welke informatie kanalen ontvang je bij voorkeur reis en evenement informatie? (meerdere antwoorden mogelijk)?

○ WhatsApp
○ Facebook
○ Instagram
○ LinkedIn
○ Twitter
○ Telegram
○ Nieuws apps of website
○ Event apps (als deze beschikbaar zijn)
○ Email
○ SMS

Deel 2 - Berichtgeving

Q17 Er volgen nu vijf berichten die je zou kunnen ontvangen op je telefoon. Voor elk bericht is de vraag:

Hoe waarschijnlijk is het dat je je eigen plan wijzigt aan de hand van dit bericht?

De schaal die daarvoor gebruikt wordt loopt van 1 (links) Helemaal niet tot 5 (rechts) Zeker.

-Het is momenteel erg druk in de bus van 14.00 uur richting het muziekfestival in Amsterdam. Reis mits mogelijk mee met een volgende event bus. Bekijk hier de locatie van de bushalte en de actuele aankomst- en vertrektijden.
-In verband met een verkeersongeluk vormt zich een file op de toegangswegen naar het muziekfestival in Amsterdam. Klik hier voor de nieuwe snelste route en voorkom files.
-Kom niet meer met de auto naar het muziekfestival in Amsterdam! Er veel vertragingen op de wegen - en de parkeerplaatsen zijn overvol. Ook elders in de stad is geen parkeergelegenheid meer beschikbaar. Wil je toch nog naar het muziekfestival in Amsterdam? Kies dan voor de trein of fiets.
-De parkeersituatie is aangepast als het muziekfestival in Amsterdam plaatsvindt. Bespaar tijd en boek je je parkeerticket vast online!
-De event bus van station Amsterdam Centraal naar het muziekfestival in Amsterdam van 15.00 uur is uitgevallen. Wij voorspellen dat er meer reizigers bij de bus stop bij Amsterdam Centraal zullen zijn dan dat er in de bus passen. De volgende bus vertrekt om 15.30 uur.

Deel 3 - Doelgroepen

De volgende vragen zijn bedoeld om een beter beeld te krijgen van de verschillende reizigers. Het onderzoek is anoniem, er zal vertrouwelijk met deze informatie om worden gegaan. Deze informatie zal op geen enkele wijze terug te leiden zijn.

Q18 Wat is je leeftijd?
Getal

Q19 Met welk geslacht identificeer jij je?

○ Man
○ Vrouw
○ Anders
○ Zeg ik liever niet

Q20 Wat is je hoogst afgeronde opleidingsniveau?

○ Basisonderwijs
○ Middelbaar onderwijs
○ Mbo, hbo-, wo-bachelor
○ Hbo-, wo-master
○ PhD of hoger
○ Geen afgeronde opleiding

Q21 Hoe bekend ben je in Amsterdam?

○ Ik woon er
○ Ik heb er gewoond
○ Ik kom er elke maand
○ Ik kom er een paar keer per jaar
○ Ik ben er wel eens geweest
○ Ik ben er nog nooit geweest

Q22 Met welke vervoersmiddelen kan je reizen naar Amsterdam? (meerdere antwoorden mogelijk)

○ Auto (eigen bezit)
○ Auto (makkelijk te lenen of abonnement voor shared car service)
○ Openbaar Vervoer
○ Auto tot P+R locatie, vanaf daar met het Openbaar Vervoer
○ Fiets, e-bike of scooter (eigen bezit of deelmobiliteit)

D

Summary tables of the results

Table D.1: Preferred travel mode per event type, related to Figure 4.2

	Music festival		Light festival		Football match	
	Frequency	Percent	Frequency	Percent	Frequency	Percent
Own car	50	13,2%	58	15,3%	90	23,8%
Shared or borrowed car	3	0,8%	3	0,8%	10	2,6%
Public Transport	251	66,4%	244	64,6%	215	56,9%
P+R	20	5,3%	21	5,6%	15	4,0%
Bicycle	54	14,3%	52	13,8%	48	12,7%
Total	378	100,0%	378	100,0%	378	100,0%

Table D.2: Preparation moment compared with preferred communication moment per event type, related to Figure 4.3

	Music festival				Light festival				Football match			
	Preperation moment		Preferred communication moment		Preperation moment		Preferred communication moment		Preperation moment		Preferred communication moment	
	Frequency	Percent	Frequency	Percent	Frequency	Percent	Frequency	Percent	Frequency	Percent	Frequency	Percent
More than a week in advance	25	6,6%	44	11,6%	25	6,6%	26	6,9%	32	8,5%	28	7,4%
A week in advance	-	-	117	31,0%	-	-	88	23,3%	-	-	88	23,3%
A few days in advance	153	40,5%	120	31,7%	112	29,6%	104	27,5%	145	38,3%	89	23,5%
One day in advance	132	34,9%	62	16,4%	115	30,4%	90	23,8%	110	29,1%	109	28,9%
A few hours in advance	38	10,1%	-	-	83	22,0%	-	-	63	16,7%	-	-
Just before departure	22	5,8%	25	6,6%	31	8,2%	48	12,7%	14	3,7%	43	11,4%
During the trip	-	-	1	0,3%	-	-	2	0,5%	-	-	2	0,5%
Leave without preparation	6	1,6%	9	2,4%	11	2,9%	20	5,3%	8	2,1%	19	5,0%
Travel company prepares the trip	2	0,5%	-	-	1	0,3%	-	-	6	1,6%	-	-
Total	378	100,0%	378	100,0%	378	100,0%	378	100,0%	378	100,0%	378	100,0%

Table D.3: Desirable travel or event information per event type, related to Figure 4.4

	Music festival		Light festival		Football match		Total
	Frequency	Percent	Frequency	Percent	Frequency	Percent	
The best departure time	122	32,3%	108	28,6%	157	41,5%	387
Route advice	218	57,7%	200	52,9%	195	51,6%	613
Crowdedness at location	107	28,3%	169	44,7%	140	37,0%	416
Crowdedness at parking	72	19,0%	72	19,0%	119	31,5%	263
Location of parking	141	37,3%	96	25,4%	133	35,2%	370
Recent traffic disruption	215	56,9%	183	48,4%	233	61,6%	631
Event schedule changes	226	59,8%	196	51,9%	176	46,6%	598
Disability information	28	7,4%	21	5,6%	26	6,9%	75
No information	6	1,6%	15	4,0%	17	4,5%	38
Total	1129	298,7%	1045	276,5%	1179	311,9%	

Table D.4: The channels that people indicated as channels on which they would like to receive information messages, related to Figure 4.6

	Frequency	Percent
WhatsApp	183	48,4%
Facebook	15	4,0%
Instagram	56	14,8%
LinkedIn	4	1,1%
Twitter	6	1,6%
Telegram	1	0,3%
News apps or websites	83	22,0%
Event apps (if available)	158	41,8%
Email	303	80,2%
SMS	104	27,5%

Table D.5: Statement ranking per cluster, related to Figure 4.9

Cluster		Free spiriters		Non-Amsterdam students		Private car lovers	
		Frequency	Percent	Frequency	Percent	Frequency	Percent
Departure time	1	6	3,3%	3	2,2%	6	9,8%
	2	24	13,2%	27	20,0%	15	24,6%
	3	69	37,9%	39	28,9%	15	24,6%
	4	72	39,6%	61	45,2%	21	34,4%
	5	11	6,0%	5	3,7%	4	6,6%
	Total	182	100,0%	135	100,0%	61	100,0%
Route	1	13	7,1%	9	6,7%	2	3,3%
	2	11	6,0%	5	3,7%	4	6,6%
	3	16	8,8%	7	5,2%	6	9,8%
	4	71	39,0%	51	37,8%	25	41,0%
	5	71	39,0%	63	46,7%	24	39,3%
	Total	182	100,0%	135	100,0%	61	100,0%
Mode of transport	1	21	11,5%	12	8,9%	5	8,2%
	2	14	7,7%	9	6,7%	8	13,1%
	3	23	12,6%	26	19,3%	14	23,0%
	4	65	35,7%	38	28,1%	21	34,4%
	5	59	32,4%	50	37,0%	13	21,3%
	Total	182	100,0%	135	100,0%	61	100,0%
Parking availability	1	21	11,5%	13	9,6%	7	11,5%
	2	26	14,3%	15	11,1%	5	8,2%
	3	33	18,1%	22	16,3%	17	27,9%
	4	64	35,2%	45	33,3%	18	29,5%
	5	38	20,9%	40	29,6%	14	23,0%
	Total	182	100,0%	135	100,0%	61	100,0%
Public transport timetable	1	14	7,7%	7	5,2%	10	16,4%
	2	37	20,3%	23	17,0%	13	21,3%
	3	48	26,4%	43	31,9%	19	31,1%
	4	67	36,8%	43	31,9%	16	26,2%
	5	16	8,8%	19	14,1%	3	4,9%
	Total	182	100,0%	135	100,0%	61	100,0

Preferences of visitors of mass events towards travel information messages

B.D. den Hollander

Delft University of Technology
April 2023

Abstract. During the influx to mass events in the past, situations have regularly arisen that people experienced as unpleasant or even unsafe. Nowadays, many researchers focus on regulating this influx by influencing travel choices, where only little attention is paid to travellers' preferences. This research identifies the preferences of different type of visitors towards receiving information, and the extent to which visitors with different information profiles act on the information messages they receive. A survey was distributed (N=378), and the number of variables was reduced by performing an Exploratory Factor Analysis. The Latent Class Cluster Analysis distinguished three clusters based on their information message preferences. The results show that people are most interested in receiving information via email a few days before departure and that it is challenging to inform them at a later time. The most impressionable cluster mainly consists of young people, unfamiliar with the event area. Event organisations can apply the results by better-aligning communication with the preferences of their visitors. The differences in behavioural intentions between the clusters turned out to be limited. This must be further investigated since it is not in line with expectations.

Keywords: Travel behaviour, information provision, message preferences, Latent Class Cluster Analysis, Exploratory Factor Analysis, behavioural intention

Introduction

Crowd formation takes place all over the world. Usually, this does not cause any problems, but sometimes situations arise where people feel uncomfortable or even endangered. A reason why so many people come together can be the influx to a mass event. A mass event is a planned gathering of a large group of people for a common purpose at a specific location [1]. If unsafe situations arise, this is often during the influx because many people want to arrive simultaneously. It is essential to spread the visitors more evenly over time and space to prevent this from happening, based on the current (or expected) level of crowdedness. To achieve this, visitors must be informed with travel and event information.

Nowadays, almost everyone owns a mobile phone, which has created new possibilities for informing people. Together with its partners, Sweco developed a tool called Crowd Management System (CROMAS). CROMAS measures the current crowdedness and predicts the prospective crowdedness based on historical and real-time data, visualised in an interactive dashboard. In case the situation does not correspond to the desired situation, measures can be taken by sending people informational messages on their mobile phones or by sending instructions to the traffic controllers [2]. This is an attempt to make people reconsider their travel choices.

In previous research, little has been discovered about the best way to approach people, especially visitors of mass events. To improve the effectiveness of CROMAS, CROMAS must be better attuned to the information preferences of different travellers. This research aimed to identify the preferences of different types of visitors towards receiving information, and to what extent visitors with different information profiles act on the information messages they receive. The focus is on the influx to mass events in Amsterdam, for which the preferences regarding the content, communication channels, and timing of the message received via mobile phones, were examined.

Visitor profiles were identified by performing a Latent Class Cluster Analysis (LCCA), in which visitors are probabilistically assigned to a cluster based on an unobserved variable [3]. After which was examined whether their behavioural intentions aligned with their stated preferences.

The remainder of this paper is structured as follows. First, the methods used, consisting of data collection and data analysis, are discussed in more detail. The data was collected by distributing a survey, the design of which is explained. This is followed by the data analysis, consisting of Exploratory

Factor Analysis (EFA), LCCA, and behavioural intention analysis. The results, including sample data, are discussed in the Results section. The final section provides this research's conclusion, implications, limitations, and recommendations.

Methodologies

Survey design

The research question consists of two parts. First, people's preferences regarding receiving information messages were examined, and after that, it was studied to what extent their preferences corresponded to their behavioural intentions. This distinction is also reflected in the survey. The first part examined the information communication preferences for three different trip purposes (leisure-related mass events in Amsterdam: Music festival (M), Light festival (L), and Football match (F)). In the second part, the behavioural intention was investigated based on five messages in which the participants indicated how likely they would act on these messages. The survey ended with part three, where some personal factors were asked. The survey was designed using Qualtrics and distributed online in Dutch.

The theory has shown that people make their modality, route, and departure time choices based on many different choice factors, which can differ per person. Only a small part of this can be influenced in the short term. The most frequently mentioned factors in the analysed studies about influencing travel choices were included in this research. These factors were expected to influence the choices since these factors showed significant effects in previous studies. In addition, factors not (often) found in previous research have been included, mainly related to the message. As a result, the event-related questions included the factors: message content, trip preparation channels, preferred time to receive information, preferred communication channels, the preferred mode of transport, and trip preparation timing. The personal factors questions included the factors: education level, age, gender, familiarity with Amsterdam, and mode of transport options.

Exploratory Factor Analysis

Three factor analyses were conducted to reduce the number of variables by combining related ones, and to gain insight into the underlying theoretical structure of the data [4]. Factor analysis is based on the idea that there are deeper underlying factors that connect the variables. The Exploratory Factor Analysis was most suitable for this research compared to the Confirmatory Factor Analysis (CFA) since no previous study was found that examined the information message preferences of visitors of mass events, so there was no hypothesis to test. In addition, the observed variable (survey question answer) could potentially be a measure of every factor. The EFA was performed in IBM SPSS Statistics (version 28). Principal Components Analysis was used as a factor extraction method for all analyses, and the orthogonal rotation with the Varimax technique was used.

First, it was checked whether the data were suitable for factor analysis. The data are suitable if the sample size is at least 300 [5], the Kaiser-Meyer-Olkin (KMO) is ≥ 0.50 (measures sample sufficiency for each variable in the model and for the complete model), and the Barlett's Test of Sphericity is significant ($p < 0.05$) (compares the correlations between the observed correlation matrix and the identity matrix) [6]. Furthermore, all the extractions of communalities have to be larger than 0.3.

The number of factors was determined based on the Kaiser rule (eigenvalues > 1) [7], which was verified by the scree plot and parallel analysis (actual values > parallel analysis values) [8].

Finally, the reliability was tested by calculating Cronbach's Alpha values (> 0.6) [9].

Latent Class Cluster Analysis

The LCCA was performed to find groups with similar message preferences. The LCCA was executed in LatentGOLD 5.0 after the number of variables was reduced by performing EFA. The LCCA statistically identifies latent classes within a population where people in the same class have similar preferences [3]. "The goal is to maximise homogeneity within clusters and heterogeneity between clusters" [10].

First, the measurement model was estimated, including only indicators which define the latent variable. In this research, the indicators were the preferred moment of reception, trip preparation moment, trip preparation channels, and preferred message content. These (message-related) factors define the latent variable, which is the attitude towards information messages. The latent variable considers the association between the indicators, which are assumed to be independent.

The measurement model was used to determine the optimal number of classes. To do so, three criteria were checked [11]. The value for both Bayesian Information Criteria (BIC) and Akaike Information Criterion (AIC) needs to be as low as possible since this indicates the most parsimonious model [12]. The third criterion concerns the Bivariate Residuals (BVR), for which a maximum value of 3.84 is allowed. Because a value above 3.84 indicates statistically sig-

nificant indicators at a 5% level. This is undesirable as it would imply a significant relationship between the indicators. If the requirements were not met, the number of indicators had to be reduced until a suitable model was found.

Second, the structural model was estimated, in which the covariates were added to the model, which describe the characteristics of the individuals in that class, also called observed variables. In this study, the covariates were the personal factors that influence the message preferences. These factors do not define the message preferences but might influence them, which is exactly what a covariate does. To start with, all covariates were added to the model. Backward elimination (elimination of the covariate with the smallest effect on the model fit) was applied until all covariates were significant. Significant covariates are indicated by Wald values higher than 3.84 and p-values below 0.05. This is desired since only a significant relation means that the covariate is likely to predict the class membership in the population [10]. The eliminated covariates were made inactive. Inactive covariates do not affect class membership but can be used to describe the classes.

Finally, the model's accuracy was investigated by checking whether the entropy R-squared value was higher than 0.8. Indicating that the LCCA most likely performed best in recovering the true value used in the simulation study, and demonstrated relatively good coverage and power in the examined settings [13].

Behavioural intention

This part of the study investigated whether the indicated preferences of the clusters aligned with their behavioural intention. The results of the second part of the survey were used for this. First, the participants were assigned to a cluster based on the highest probability, leading to "hard partitioning" [14]. After this, the distributions of the scores among the clusters were compared. This gave a first indication of the actual follow-up behaviour. Since only five messages were shown to the participants, no firm conclusions could be drawn, but striking observations may lead to future research.

Results

Descriptive personal results

The survey resulted in a sample size of 378 (response rate 70%). The participants represent the Dutch population quite well based on gender but not so well based on age and level of education [15]. The participants are significantly younger and higher educated compared to the Dutch population. Since the average event visitor is younger than the average Dutch inhabitant, the actual differences are smaller, so the biggest difference between the participants and the Dutch population is based on education level.

Descriptive data results

The general results show a lot of overlap between the different events. The preparation moment distribution is compared with the preferred communication moment distribution in Figure 1. People would like to receive information for all three events before preparing their trip, preferably a few days before the event takes place (a maximum of one week in advance). This can be seen in Figure 1 because the blocks on the left are relatively larger than the blocks on the right side for COM compared with PREP per event. Everyone must receive this information, as hardly anyone leaves the entire preparation to their travel company.

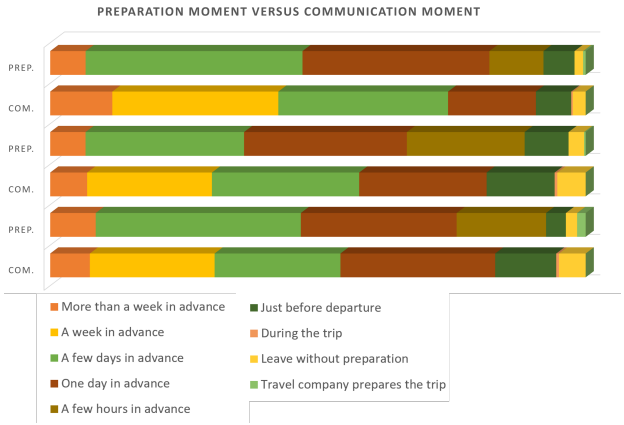

Figure 1: *Preparation moment compared with preferred communication moment per event type (PREP. = preparation moment and COM. = communication moment)*

It is important to provide visitors only with the necessary information and avoid bombarding them with additional information to increase the likelihood of them taking action on the information provided. The desired content is illustrated in Figure 2, with participants being allowed to select multiple answers. Each participant selected an average of three types of information, with recent traffic disruptions, route advice, and event schedule changes as the most popular content.

Because visitors tend to look up less information en route or at the location, as much information as possible (including information about the return trip) must be communicated before departure. This can be seen from the results in Table 1 by looking at "total sources consulted", which are much higher before

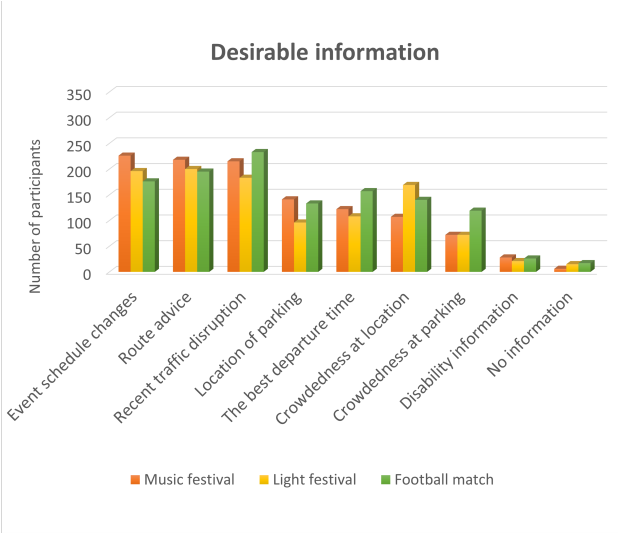

Figure 2: *Desirable travel or event information per event type*

departure than en route, and at the location. These findings are important to take into account if visitors are to be informed. Since this data show many more possibilities for influencing before departure, the rest of this study focuses on before-departure data. Before departure, the organisation's website, emails, and navigation systems are the most consulted channels, and social media are not very popular.

Table 1: *Preparation channels per event type and moment, where M = Music festival, L = Light festival, F= Football match*

	Before departure			En route			At location		
	M	L	F	M	L	F	M	L	F
The organisation's website	303	309	236	46	48	35	49	60	26
Event organiser emails	268	196	218	38	22	36	31	15	19
SM channels of the organisation	127	118	98	47	47	41	43	40	34
SM channels of other organisations	44	47	39	26	21	23	19	19	15
organisation's app (if available)	95	97	91	64	51	59	131	108	77
Navigation systems such as Google Maps	207	193	225	257	231	227	66	101	68
Total sources consulted	1044	960	907	478	420	421	339	343	239

To inform people, it is important to know which channels people use to prepare their trip and through which channels they prefer to receive information messages. The results are visualised in Figure 3, in which participants were allowed to give multiple answers. The most notable value is for email, which is considerably higher than the other values. Other high values can be seen for WhatsApp, event apps, and SMS. Even though event organisations increasingly focus on social media, these are not the channels through which people like to receive information.

Exploratory Factor Analysis

The EFA was performed trice, in which no variables had to be deleted to find a suitable result. The criteria were checked, with the result that the 6 variables

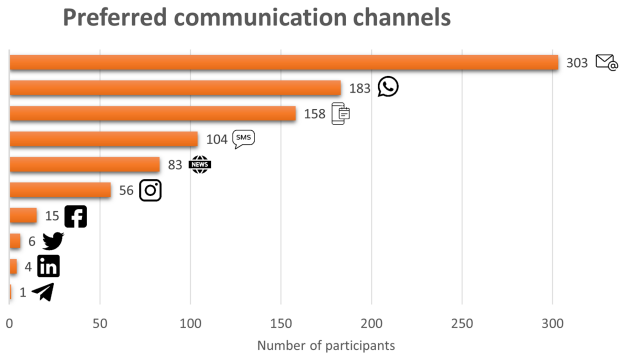

Figure 3: *Preferred communication channels*

related to the moment of communication were combined in one factor, the 24 variables related to content were combined in 7 factors, and the 18 variables related to communication channels were combined in 4 factors. The variables for different events but with the same topic were mostly aggregated. In addition, some topics were merged, such as the organisation's and third parties' social media channels. The factors were used as indicators in the LCCA.

Latent Class Cluster Analysis

A suitable model was found with seven indicators, including six content-related indicators (parking, traffic disruption, route advice, departure time advice, event schedule, and level of crowdedness) and one indicator about timing. Backward elimination started with 22 active covariates and resulted in a significant model, including 13 active covariates (age, gender, familiarity with Amsterdam, 4 travel mode options (own car, shared car, public transport, bicycle), 4 preferred communication channels (WhatsApp, Instagram, news, SMS), and preferred travel mode to music and light festival), and 9 inactive covariates (level of education, 1 travel mode option (P+R), 6 preferred communication channels (Facebook, LinkedIn, Twitter, Telegram, Event Apps, Email), and preferred travel mode to football match).

The LCCA resulted in three clusters, each with its own information preferences and personal characteristics. The results can be seen in Table 2. Furthermore, the information provision preferences of the average person and the three clusters are visualised in Figure 4.

The average person is 37 years old, not more often male or female, highly educated (52% HBO, WO master or higher), visits Amsterdam a few times a year or more often, and has the option to travel to Amsterdam by public transport (94%), by their own car (61%), or via a P+R location (51%). The average person prefers to receive messages about recent

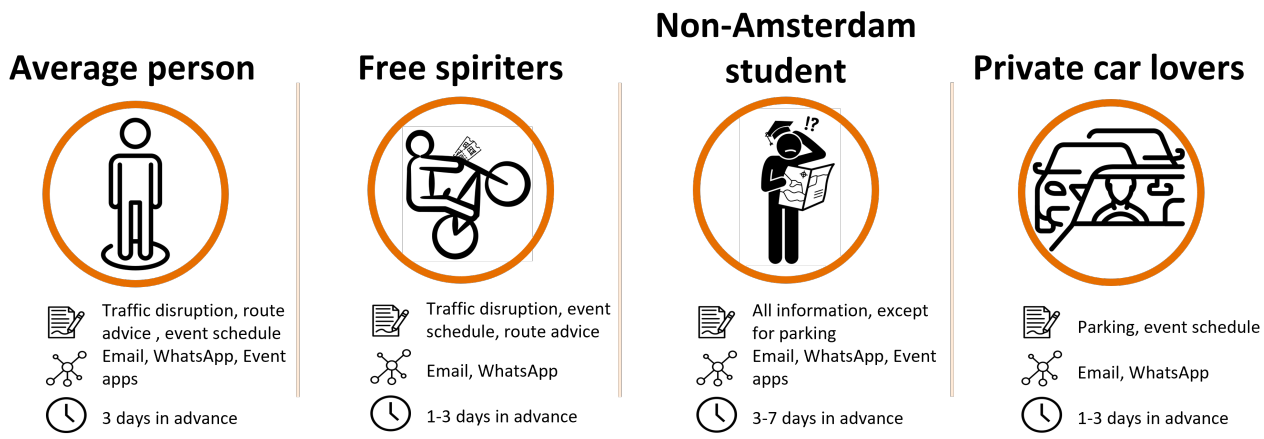

Figure 4: *Information provision preferences per cluster*

traffic disruptions (indicated by 56%), route advice (indicated by 54%), and/or event schedule changes (indicated by 53%). Ideally, the average person receives these messages via email (indicated by 80%) and otherwise through WhatsApp messages (indicated by 48%) before trip preparation. For 93% of the participants, this means one day in advance or more, but with a maximum of one week in advance.

The extent to which each cluster is interested in receiving certain information is shown in Figure 5. The values indicate how often someone indicated to be interested in that type of information. The participants could answer this question with yes or no for all three events. The meaning of the numbers is as follows:

#1	Parking information
#2	Traffic disruption
#3	Route advice
#4	Departure time
#5	Event schedule changes
#6	Crowdedness

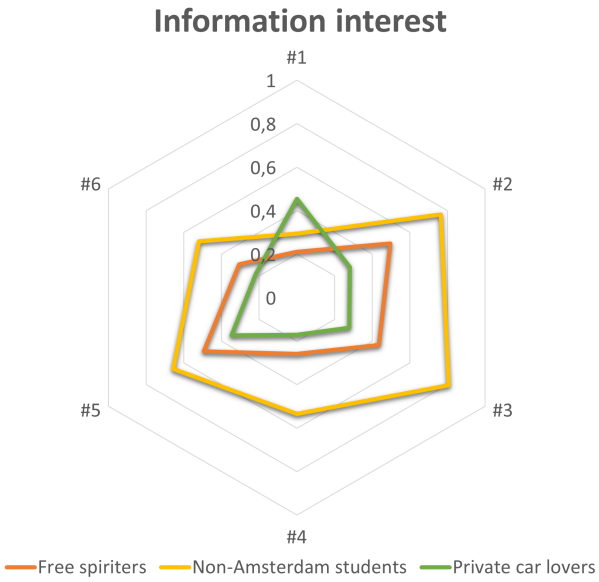

Figure 5: *Interest in certain information content per cluster, frequency on the range 0 (never) to 1 (always)*

The free spiriters (49% of the participants) have a less pronounced preference for a mode of transport than the other clusters, but prefer travelling by public transport (64%) or bicycle (20%), and mostly not the car (7%). They are relatively familiar with Amsterdam, mainly men (62%), and they are highly educated (56% HBO, WO master or higher). They would like to receive information messages about recent traffic disruptions, event schedule changes, and/or route advice by email between 1 and 3 days in advance.

The non-Amsterdam students (35% of the participants) are unfamiliar with Amsterdam, and they are still very young (almost half of the group is between 15 and 25 years old). Therefore, they often have not finished school yet, meaning they are less educated. They crave information and are not very critical of the content and communication channels. However, they prefer to receive messages containing route advice by email between 3 and 7 days in advance. This cluster is the most impressionable cluster.

The private car lovers (16% of the participants) are mainly men (61%) who regularly visit Amsterdam, and they are creatures of habit. They almost all own a car (95%) and really prefer to use it to drive to the events compared with the other clusters (64% compared with 7% and 11%). Because they have their habits, they are not very interested in information, except parking information. They prefer to receive this information by email between 1 and 3 days in advance. This cluster is the least impressionable cluster.

Table 2: *Cluster profiles, including indicators, active covariates, and inactive covariates*

Cluster name		**Free spiriters**	**Non-Amsterdam students**	**Private car lovers**
Cluster size		49%	35%	16%
INDICATORS (MEAN)				
	#1 Parking information	0,2099	0,2949	0,4548
	#2 Traffic disruption	0,4972	0,7645	0,2814
	#3 Route advice	0,4356	0,8062	0,2785
	#4 Departure time	0,259	0,5338	0,1703
	#5 Event schedule changes	0,494	0,656	0,3473
	#6 Crowdedness	0,306	0,5194	0,218
	Moment	3,3283	2,6747	3,5211
ACTIVE COVARIATES				
Age	15-25	29,2%	47%	23%
	26-35	22,2%	24%	25%
	36-45	8,1%	2%	15%
	46-55	17,3%	16%	23%
	56-65	19,5%	10%	13%
	66+	2,7%	2%	2%
	Mean	39	33	39
Gender	Male	62%	34%	61%
	Female	38%	65%	36%
	Other	0%	0%	3%
	Blank	0%	1%	0%
Familiarity with Amsterdam	I live there	10%	0%	3%
	I lived there	11%	2%	10%
	Visit once a month	18%	7%	18%
	Visit few times a year	49%	59%	44%
	Been few times	12%	32%	23%
	Never been	1%	0%	2%
Possible to go to Amsterdam by	Own car	57%	53%	95%
	Shared car	12%	33%	30%
	Public Transport	95%	99%	80%
	Bicycle	20%	5%	23%
Prefer receiving information via	WhatsApp	44%	58%	42%
	Instagram	5%	28%	17%
	News	14%	32%	25%
	SMS	26%	37%	9%
Preferred transport mode M	Own car	4%	6%	57%
	Shared car	0%	1%	3%
	PT	64%	85%	35%
	P+R	10%	1%	2%
	Bicycle	23%	8%	3%
Preferred transport mode L	Own car	2%	7%	75%
	Shared car	1%	2%	0%
	PT	67%	83%	19%
	P+R	10%	1%	2%
	Bicycle	21%	8%	5%
INACTIVE COVARIATES				
Education	Primary education	0%	1%	0%
	Secondary education	9%	17%	2%
	MBO, HBO, WO bachelor	35%	39%	38%
	HBO, WO master	51%	42%	57%
	PhD or higher	5%	2%	3%
	Not educated/unknown	0%	1%	0%
Possible to go to Amsterdam by	P+R	44%	55%	65%
Prefer receiving information via	Facebook	2%	5%	10%
	LinkedIn	1%	0%	5%
	Twitter	2%	1%	3%
	Telegram	1%	0%	0%
	Event Apps	37%	52%	33%
	Email	80%	86%	68%
Preferred transport mode F	Own car	15%	20%	59%
	Shared car	1%	6%	2%
	PT	61%	63%	30%
	P+R	7%	1%	0%
	Bicycle	16%	9%	10%

Behavioural intention

The participants indicated how likely they will act on this information for five information messages. The results of this can be seen per cluster in Figure 6. The results show an unexpected amount of overlap between the different clusters. Although the clusters showed different information preferences, the differences between these results remain minimal. It seems that people did not judge the messages with the idea of travelling by a particular mode of transportation or from a specific 'cluster type view'. However, a significant difference can be seen in the message about the public transport timetable between the *private car lovers* compared with the *free spiriters*, and the *non-Amsterdam students*.

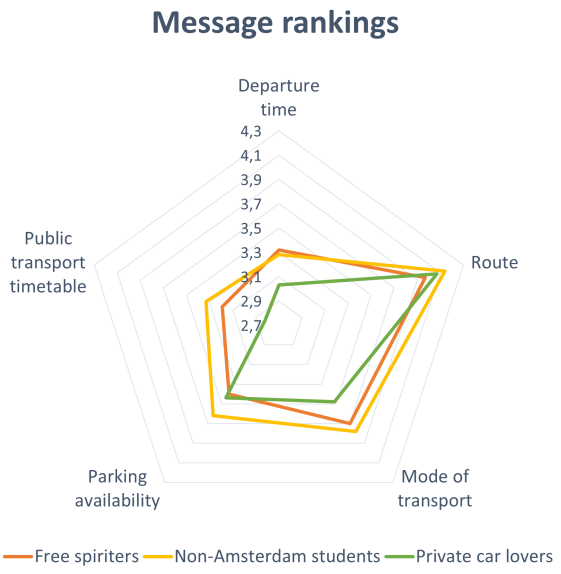

Figure 6: *Message rankings per cluster, scale ranges from 1 (absolutely not) to 5 (definitely)*

No firm conclusions can be drawn from these results because only five messages with different topics were shown to the participants. To draw firm conclusions, multiple messages on the same topic must be presented to the participants to see the influence of other aspects. The results of this research provide a first indication.

Conclusions and recommendations

Conclusion

In recent years, more research has been done into influencing travel behaviour. However, little attention was paid to the needs of the travellers. This research has contributed to this by identifying the preferences of different types of visitors towards receiving information and the extent to which visitors with different information profiles act on the information messages they receive.

The general results clearly showed that sharing as much information as possible with people before departure is essential. In addition, event organisations are increasingly focusing on social media, but the participants expressed a clear preference for communication via email. The results also showed that the information communication preferences for the three included events differed only a little.

The LCCA distinguished three clusters based on preferences towards receiving information. The non-Amsterdam students turned out to be the most impressionable cluster. In addition, they are very interested in receiving a lot of information, which directly offers the opportunity to influence their choices. The private car lovers are least impressionable, except when it comes to car-related information. The free spiriters are familiar with Amsterdam, which makes them particularly interested in information about changes. The behavioural intention in this study did not correspond much with the differences in information preferences of the clusters, which showed that further research is needed to find out what caused these results.

Implications

In this study, clusters emerged that were not distinguished in this way before. The existence of these clusters and the apparent effect of different aspects of the messages on information preferences appear to be important elements to keep in mind in future research on travel behaviour.

The results of this research are relatively easy to implement in practice. Event organisations and Sweco can better tailor communication with visitors to the target group's preferences. This does not require large investments, which makes short-term implementation possible. The most important aspects are that as much information as possible is to be sent via email a few days before the event and not via social media, where it is most important that the non-Amsterdam students receive the information that interests them.

Limitations

The survey was distributed online among a convenience sample. As a result, the sample is higher educated than the average Dutch population. It has not become clear from this study whether a new cluster will emerge if the population of the Netherlands is better represented or whether these people will spread over the existing clusters. Second, there

seem to have been interpretational differences when assessing the messages in the survey. This misunderstanding could have been prevented by telling the participants, before starting with this second part of the survey, that they had to judge the messages from the perspective that they were travelling by their stated preferred mode. In addition, the connection between the information preferences and the behavioural intention would have been more complete if the timing and communication channel would also have been specified, instead of the focus being on the content only.

Recommendations

Several recommendations for follow-up research emerged from this study. The first recommendation for follow-up research resulted from the limited differences in behavioural intention between clusters after seeing the information messages. It is expected that there are more significant differences between the clusters than what was found in this study, caused by the perspective from which people assessed the message. It is important to investigate this more extensively to understand the importance of information preferences.

The second aspect that needs more attention is the way information is presented. This was not often included in previous studies. Still, based on the results of this research, it is expected that more message-related factors than content, timing, and communication channels affect travel choices. Learning more about these aspects is essential to get the best effect.

Thirdly, the results showed that people are most interested in receiving information before departure. However, if important information still needs to be communicated later, it must be investigated how people can best be reached when they are en route or already at the location. It is still hard to inform people at later moments since they look up less information and are therefore less easily accessible.

References

[1] J. M. F. Rodrigues. *What is Mass Events | IGI Global*. Nov. 2017. URL: https://www.igi-global.com/dictionary/mass-events/61900.

[2] CROMAS consortium. *Cromas Consortium*. 2022. URL: https://cromas.io/.

[3] Linda M. Collins and Stephanie T. Lanza. *Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences*. Vol. 718. John Wiley & Sons, 2009.

[4] Statistics Solutions. *Factor Analysis*. Aug. 2021. URL: https://www.statisticssolutions.com/free-resources/directory-of-statistical-analyses/factor-analysis/.

[5] Barbara G Tabachnick, Linda S Fidell, and Jodie B Ullman. *Using multivariate statistics*. Vol. 6. pearson Boston, MA, 2013.

[6] Brett Williams, Andrys Onsman, and Ted Brown. "Exploratory factor analysis: A five-step guide for novices". In: *Australasian journal of paramedicine* 8 (2010), pp. 1–13.

[7] A. Field. *Discovering Statistics Using SPSS*. ISM (London, England). SAGE Publications, 2009. ISBN: 9781847879073. URL: https://books.google.nl/books?id=a6FLF1YOqtsC.

[8] Bruce Thompson. "Exploratory and confirmatory factor analysis: Understanding concepts and applications". In: *Washington, DC* 10694.000 (2004).

[9] Khairul Azhar Mat Daud et al. "Validity and reliability of instrument to measure social media skills among small and medium entrepreneurs at Pengkalan Datu River". In: *International Journal of Development and sustainability* 7.3 (2018), pp. 1026–1037.

[10] Maarten Kroesen. *SEN1721 module 3: Latent class cluster models introduction*. 2022.

[11] Eric Molin, Patricia Mokhtarian, and Maarten Kroesen. "Multimodal travel groups and attitudes: A latent class cluster analysis of Dutch travelers". In: *Transportation Research Part A: Policy and Practice* 83 (2016), pp. 14–29.

[12] Jeroen K Vermunt and Jay Magidson. "Latent GOLD® choice 4.0 user's manual". In: *Statistical Innovations Inc., Belmont, MA* (2005).

[13] Christine DiStefano and RW Kamphaus. "Investigating subtypes of child development: A comparison of cluster analysis and latent class cluster analysis in typology creation". In: *Educational and Psychological Measurement* 66.5 (2006), pp. 778–794.

[14] Zsuzsa Bakk, Daniel L Oberski, and Jeroen K Vermunt. "Relating Latent Class Assignments to External Variables: Standard Errors for Correct Inference." In: *Political analysis* 22.4 (2014).

[15] CBS. *Bevolkingspiramide*. 2023. URL: https://www.cbs.nl/nl-nl/visualisaties/dashboard-bevolking/bevolkingspiramide.