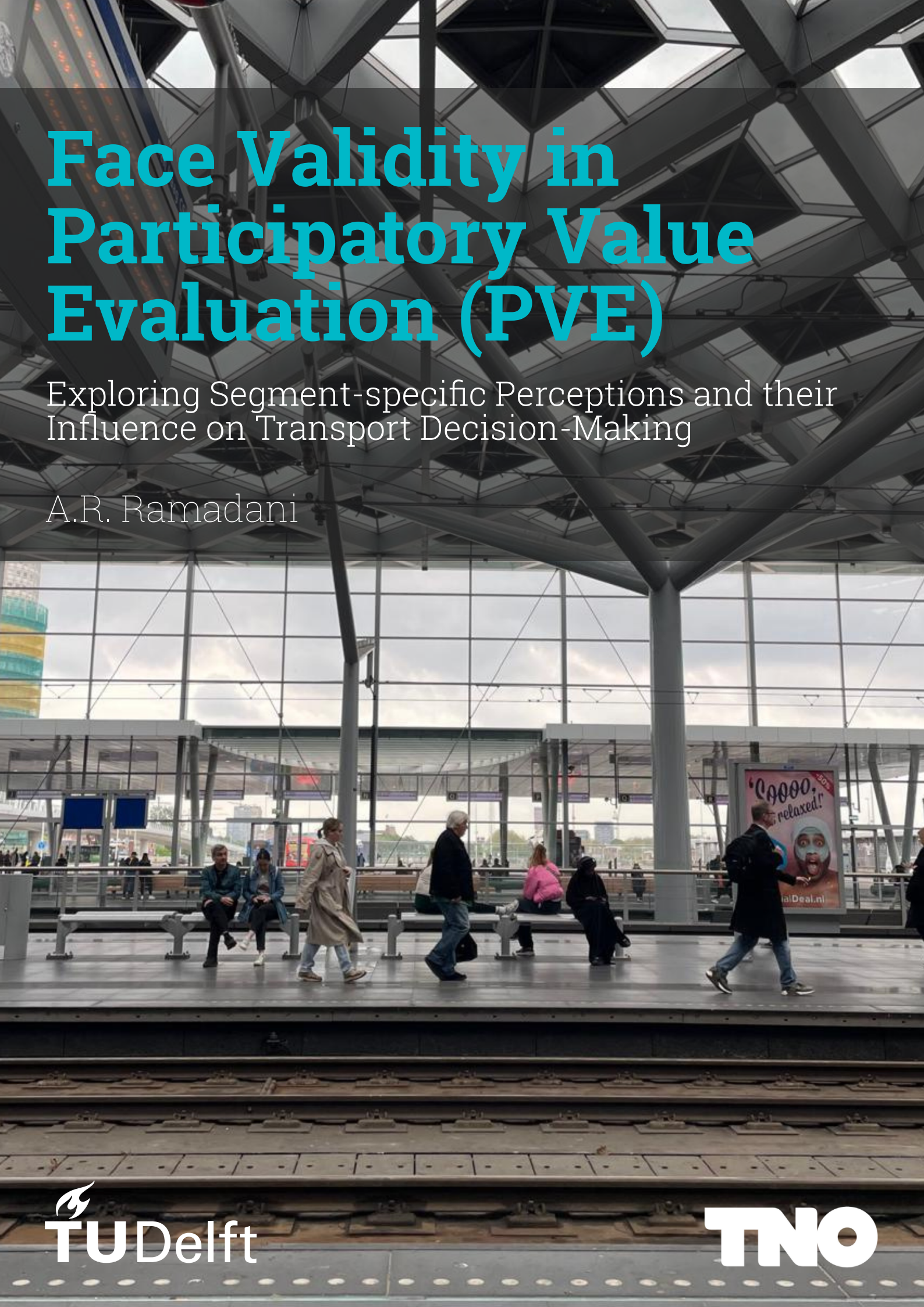
Face Validity in Participatory Value Evaluation (PVE)

Exploring Segment-specific Perceptions and their Influence on Transport Decision-Making

A.R. Ramadani



TU Delft

TNO

Face Validity in Participatory Value Evaluation (PVE)

Exploring Segment-specific Perceptions and their Influence on Transport Decision-Making

by

A.R. Ramadani

in partial fulfilment of the requirements for the degree of

Master of Science

in Transport, Infrastructure, and Logistics

at the Delft University of Technology.

to be defended publicly on August 27th 2024.

Student Number:	5347734	
Project Duration:	February, 2024 - August, 2024	
Thesis Committee:	Dr. Mr. N. (Niek) Mouter	TU Delft, chair
	Dr Ir. N. (Niels) van Oort	TU Delft, supervisor
	A.M. (Anna) de Ruijter	TU Delft, advisor
	Ir. E.S. (Eric) Tol	TNO, external supervisor

TU Delft

TNO

Preface

This thesis marks the end of a two-year journey in the Transport, Infrastructure, and Logistics program—a journey that was once just a distant dream, some 6 to 8 years ago. Back then, the idea of pursuing a master's degree seemed almost impossible. But here I am, having overcome countless moments of self-doubt and long periods of contemplation before finally deciding to study far from home in a field quite different from my previous work. The challenges were many, but so were the lessons. Now, as I reach the finish line, I find myself deeply grateful—not only to my God but also to the person who made this happen: myself. I hope this thesis and the journey it represents will always remind my future self of the power of courage and perseverance.

This work wouldn't have been possible without the guidance and continuous feedback from my thesis committee. First and foremost, I want to thank Niek for introducing me to the novel method of Participatory Value Evaluation (PVE) and allowing me to explore it as my thesis topic. Early in the process, you gave me a piece of advice that has stuck with me: "Don't create unnecessary problems." It's advice I'll carry with me far beyond academia. I'm also grateful to Niels, whose guidance and insistence on deeply reflecting on the implications of my results sharpened my critical thinking skills. A special thank you to Anna, who played a crucial role in the early stages of my thesis, guiding me through data processing and introducing me to the concept of face validity.

I'm also deeply thankful to Eric, Eleni, and Maiara, for giving me the opportunity to undertake my thesis internship at TNO. Thank you for introducing me to the concept of Broad Welfare and for encouraging me to be confident and unafraid of making mistakes—an invaluable lesson for any student who might feel pressured to have all the answers. The cover of this thesis report features a photo taken at Den Haag Centraal, a place that holds special meaning as it reminds me of the days spent at the TNO office nearby.

I extend my heartfelt thanks to the friends I made during my studies. To Sobitil Indo, PPI Delft, RH Squads, TIL friends, my best friends back home and here, Rezzy, Fahmi, Saras, and so many others I can't list here individually—your support has been a cornerstone of my journey.

Finally, I want to thank my family. To my niece, Afiya—seeing your new pictures every day brought me so much joy and motivation to keep pushing forward with my thesis. To my sisters and brother-in-law, thank you for brightening my days with your funny stickers and random chats. To Dennis, your unwavering belief in me, even when I doubted myself, has meant the world. To my mom, who has given her all to shape the person I am today, I owe you everything. And to my late dad, I fulfilled the promise I made to you—I completed my master's degree in the Netherlands. This achievement is as much yours as it is mine.

A.R. Ramadani
Delft, August 2024

Summary

Introduction

Transportation plays a crucial role in modern society by enabling the movement of people, commodities, and animals for various purposes, including work, trade, and social interactions (Van Wee et al., 2013). Efficient transportation planning is essential for maximizing the advantages of transportation while minimizing its negative effects, such as environmental pollution and safety issues. The primary goal of transportation planning is to ensure the efficient and sustainable movement of people and goods, while improving safety, speed, comfort, convenience, and cost-effectiveness (Meyer, 2016).

In recent decades, there has been a notable shift in transportation planning toward more inclusive and deliberative approaches. The purpose of this change is to enhance the integration of public input, thus ensuring that policies are not only efficient but also widely embraced by citizens (Bickerstaff & Walker, 2001). Participatory methods, such as Participatory Value Evaluation (PVE), have become promising instruments for improving public participation in transportation planning. PVE facilitates citizen participation in the decision-making process by enabling people to evaluate and provide their input on various policy alternatives.

Despite its potential, PVE is still a relatively new method, and its validity has not been thoroughly investigated (Tuit, 2022). Understanding the validity of PVE, specifically face validity and different group of citizens evaluations on it, is essential for ensuring that the method is viewed as genuine and effective by participants. The main objective of this thesis is to fill this research gap by examining the face validity of PVE in transportation planning. Specifically, it seeks to answer the question:

”To what extent are there distinct segments in the population who (in)consistently evaluate the face validity of Participatory Value Evaluation (PVE) in transportation planning?”

Current State of Knowledge on Face Validity in Public Participation Instruments

The concept of validity is crucial in the development and assessment of a measurement tool. Validity is the extent to which the data gathered by an instrument precisely addresses the particular subject under examination (Ghauri et al., 2020). It is important to ensure the validity of citizen participation methods, such as PVE, since any perceived lack of validity can greatly affect the method’s credibility, appeal, and acceptability (de Ruijter, 2022).

Face validity refers to the extent to which a test seems to measure what it claims to measure (Tuit, 2022). This type of validity is essential for ensuring that the tool is considered relevant, easy to understand, and acceptable by its target audience. When participants see an instrument as having face validity, they are more likely to actively participate and provide reliable responses. That makes face validity often employed in assessing tools in various sectors. In the context of public participation approaches, recently face validity has been applied to evaluate the clarity, relevance, and comprehensiblity of PVE as perceived by its users.

Case Studies

This study utilizes three datasets from PVE consultations in the Netherlands, namely Lelylijn, Oude Lijn, and Mobility Vision, to analyze the face validity of PVE consultation. These projects are selected due to their relevance to transportation planning and the availability of large datasets from approximately 4,000 respondents per consultation. The Lelylijn project aims to improve transportation infrastructure in the Northern Netherlands, the Oude Lijn project focuses on enhancing a historical railway route, and the Mobility Vision project is part of a strategic plan for future transportation in the Netherlands. The descriptive results in the later stage of this study show that all three datasets yield comparable results, indicating these three datasets are suitable for comparisons in this study.

Methods

To answer the main research question, this study incorporates both quantitative and qualitative methods. Quantitative methods include descriptive analysis, crosstabulation analysis, latent class cluster analysis, and multinomial logistic regression. while qualitative methods involve content analysis and focus group meeting with experts.

Results

The findings showed that respondents' evaluations of the overall PVE consultations were generally positive, with higher average ratings compared to past PVE consultations. The average ratings were 7,42 for Lelylijn, 7,72 for Oude Lijn, and 7,55 for Mobility Vision consultations. In terms of face validity evaluation, participants expressed trust in the credibility of the consultations, although there was potential to make the PVE consultation less steering and easier to understand.

The analysis revealed that younger individuals generally rate PVE consultations more positively, men tend to give higher ratings but are more critical in certain face validity dimensions, and highly educated respondents are more critical in assessing the consultation's comprehensiveness and relevance. Frequent users of the train and those living closer to the project site also showed more positive evaluations.

Latent Class Cluster Analysis (LCCA) identified four clusters of respondents in each PVE consultation. Positive evaluators generally fall into the category of young adults with a moderate to high level of education, living close to the project and regularly using the train. Negative evaluators are typically young to middle-aged, with varying education levels, living far from the projects, and rarely using the train.

Content analysis of the comments from the Lelylijn consultation revealed that the majority of feedback came from individuals with higher education levels, predominantly men. The main concerns raised included the complexity of the consultation process, the perceived steering nature of the consultations, and the lack of comprehensive information provided. These findings suggest that while overall feedback is positive, there are specific areas that require improvement to enhance participant satisfaction and engagement. Furthermore, the focus group meeting with PVE experts also align with the findings, underscoring the necessity of improving several aspects of PVE.

Discussion and Conclusion

The primary goal of this study was to evaluate the face validity of PVE consultations in transportation planning and to identify distinct segments of the population that consistently evaluate PVE's face validity positively or negatively. The findings provide a nuanced understanding of how different demographic groups perceive PVE. Additionally, the results highlight the significant association of the respondents' characteristics such as age, gender, education level, frequency of train use, and proximity to the project site with their perceptions of PVE.

This study contributes to the field of public participation in transportation planning by providing insights into the validity of the PVE method and highlighting the importance of face validity in ensuring its effectiveness and acceptance. The findings support the suitability of PVE as a tool for more inclusive and forward-looking policy-making. Recommendations for enhancing the face validity of PVE consultations include making the consultations less complex, providing comprehensive information, and ensuring the consultation process is perceived as fair and not steering. Further research should continue to explore the demographic factors influencing face validity perceptions and investigate other validity types to strengthen PVE as a public participation tool. Further analysis conducted at the final stage of this study about the broad welfare criteria also confirms that PVE is an appropriate approach for evaluating transport-related planning tools from a broad welfare perspective.

Contents

List of Figures

List of Tables

1 Introduction

Transportation serves as a fundamental pillar of modern society, facilitating the movement of individuals, goods, and animals for various purposes such as employment, commerce, and social interactions (Van Wee et al., 2013). A diverse range of transportation modes including land (road and rail), air, water, cable, pipeline, and space exist to support these movements, underlining the critical role of transportation infrastructure in societal functionality. This infrastructure encompasses roads, railways, airways, pipelines, as well as transportation hubs like railway stations, airports, and seaports. However, alongside the evident benefits, transportation activities also bring about consequences such as environmental pollution and safety concerns for users. Hence, effective transportation planning is important to enhance the benefits of transport while mitigating its consequences. The overarching objective of transportation planning is to ensure safer, faster, more comfortable, convenient, cost-effective, and environmentally friendly movement of people, goods, and animals from one point to another (Meyer, 2016).

Transportation planning practitioners are involved in a wide range of activities, ranging from the development of transport plans, policies, and projects to their day-to-day implementation. Cascetta (2009) describes these activities, spanning short-term operational management programs, short/medium-term tactical planning, feasibility studies, and strategic/long-term planning. These activities may involve individual efforts or collaboration with other stakeholders. Historically, transportation projects have been organized around a systematic process involving goal and objective setting, problem identification, generation of alternatives, and evaluation. To enhance decision-making processes and address societal and environmental needs comprehensively, transportation planning practitioners utilize diverse tools, notably for researching and assessing the future impacts of policy interventions. One well-known example of such tools is the (Social) Cost-Benefit Analysis (SCBA), which seeks to derive a comprehensive indicator of costs and benefits for all affected stakeholders (Van Wee et al., 2013).

In recent decades, there has been a notable shift in government approaches toward understanding needs and values of citizens, moving away from one-way decision-making strategies toward more inclusive and deliberative processes. This transition is driven by the recognition that traditional approaches often overlook the perspectives of residents and hinder public participation (Bickerstaff et al., 2002). Emerging studies advocate for more deliberative approaches, highlighting their potential to enhance social cohesion, fairness, service quality, and societal learning (Ansari & Andersson, 2011). The use of participatory approaches leads to more legitimate policy decisions (Buijs & Boonstra, 2020). This transformation is reflected not only in transportation planning practices but also in other domains involving public decision-making. Consequently, legislative frameworks, especially in democratic countries, have evolved to support principles of democracy and inclusiveness in transportation planning (Cascetta & Pagliara, 2013).

Delgado et al. (2011) identified three rationales for public participation: instrumental, substantive, and normative. The instrumental rationale views citizen participation as a method to accomplish a particular objective or outcome. Participation is seen as a means to achieve specific goals, such as strengthening public perception or bolstering decision legitimacy. The substantive rationale highlights that public involvement can improve results by bringing new information to light. Citizens' participation in decision-making can provide useful insights, views, and facts that enhance or modify the decisions being made. The normative rationale is based on the premise that citizen engagement is inherently valued as it corresponds with democratic values and principles. Participation is deemed appropriate according to normative ideals to advocate for democracy, transparency, and civic involvement.

In transportation planning, citizen involvement can be realized through various means, including digital participatory approaches facilitated by advancements in technology. These approaches enable effective communication, empowerment of citizens, and cost reduction in policymaking processes (Zheng & Schachter, 2016). Among these methods, Participatory Value Evaluation (PVE) emerges as a new

promising approach. This chapter explores the PVE methodology, its integration within broader public participation frameworks, and PVE face validity. Additionally, a reading guide is provided to navigate the following discussion.

1.1 Participatory Value Evaluation (PVE)

PVE is an online experiment that enables citizens to experience the decision-making process from the perspective of policymakers. Unlike the conventional discrete choice framework, the use of PVE allows respondents to choose more than one option instead of being restricted to selecting only a single option (Bahamonde-Birke & Mouter, 2024). The participants are provided with specific information regarding the policy options, including their characteristics and potential effects. In addition, participants must take into account specific factors such as government budget constraints or sustainability objectives when choosing the policy options, which presents them with trade-offs due to their restricted budget allocation. The budget preferences supplied by respondents represent the value they assign to the options. These preferences therefore reflect both their personal utility maximization and societal preferences (Mouter et al., 2021).

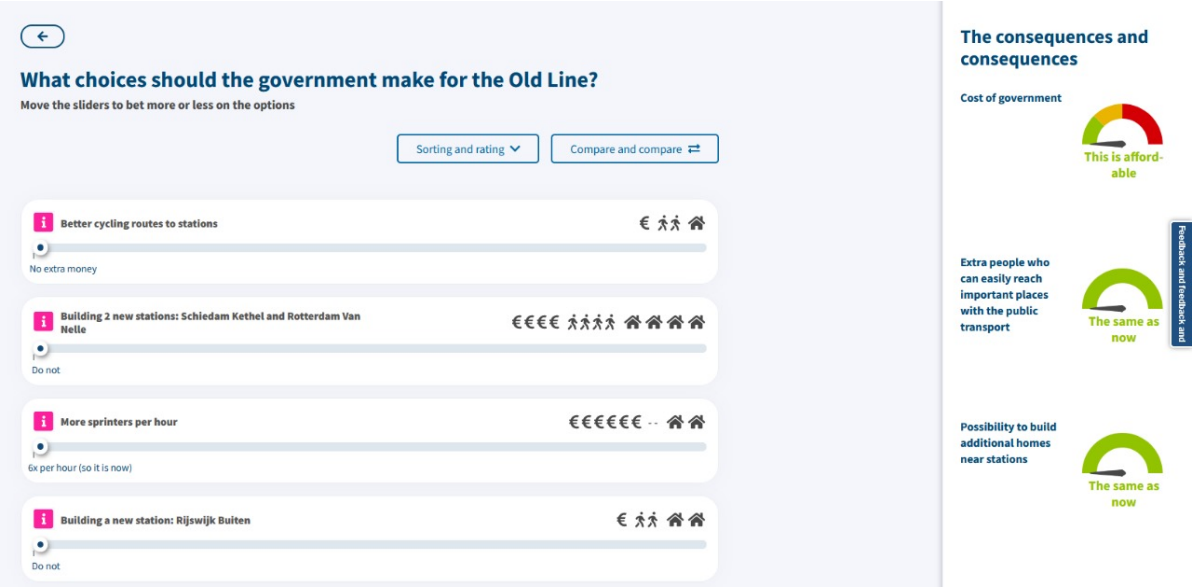


Figure 1.1: PVE example: slider choice task, to indicate preferences in Oude (Old) Lijn Consultation

The establishment of a PVE involves a structured five-stage process, as outlined by Bouwmeester (2021). Firstly, the collective policy problem requiring resolution is identified. Then, policies aimed at addressing this problem are defined and their potential impacts are established to help participants understand the consequences of each policy. Simultaneously, participants are required to adhere to constraints such as the public budget and available space limits, as illustrated in Figure 1.1, during the PVE consultation. The second and third steps then required data collection form literature search of feasibility studies and expert interviews. Following this, the PVE questionnaire is designed to reflect the information desired by policymakers and subsequently distributed for online completion by citizens. Lastly, responses from the PVE are analyzed and insights derived from these results are presented to policymakers for further decision-making.

The PVE methodology is well-aligned with assessing public policy based on broad welfare, as it captures citizens' preferences regarding the broader objectives of transportation planning. By encompassing not only the monetary effects of government projects but also non-monetary factors such as environmental considerations and social justice, PVE provides a comprehensive perspective on policy impacts (Bahamonde-Birke & Mouter, 2024; Rijkwaterstaat, 2022). This approach enables different demographic groups to share their subjective experiences of prosperity, ensuring that all essential values, preferences, and concerns are taken into account during policy deliberations (Rijkwaterstaat, 2022). By allowing citizens to express values, preferences, and concerns not always captured in traditional

decision-making processes, PVE enhances the inclusivity and legitimacy of transportation planning initiatives.

1.1.1 PVE and Public Participation

Participatory Value Evaluation (PVE) fosters two-ways information exchange between facilitators and citizens. By equipping participants with information about policy options, their properties, and constraints, PVE enhances awareness of the dilemmas faced by policymakers and the trade-offs inherent in allocating limited resources. Concurrently, facilitators can gain valuable insights from citizens regarding their choices and underlying values. Thus, following Rowe and Frewer (2004) framework on forms of participation, PVE qualifies as a form of public participation.

PVE offers a concrete method for integrating public participation into decision-making processes. Positioned within Arnstein (1969) participation ladders, PVE corresponds to the levels of informing and consulting citizens. In practice, PVE is typically initiated prior to final policy decisions. While citizens engage in selecting policy options, they can also participate in subsequent phases to provide further input and clarification of results. Ultimately, the objective is to generate recommendations for policymakers based on the insights gained through the PVE process (Spruit & Mouter, 2020).

1.1.2 PVE and Face Validity

Public participation's effectiveness necessitates a thorough assessment of its validity (Rosener, 1978; Rowe & Frewer, 2004). Validity, which gauges the extent to which an instrument measures its intended construct, is multifaceted and context-dependent (Boxebeld et al., 2024; Gaber & Gaber, 2010; Tuit, 2022). Despite its importance, validity often remains unexamined in commonly used instruments due to various constraints (Bannigan & Watson, 2009; Patel & Desai, 2020). However, neglecting validity checks, particularly in instruments like PVE consultations, can lead to misleading results and misguided decisions (Anastasi & Urbina, 2007).

To assess validity, multiple validation procedures are recommended, with face validity being one of the quickest methods (Bannigan & Watson, 2009; Gaber & Gaber, 2010; Tuit, 2022). Face validity evaluates whether, on the surface, the instrument appears to make sense and is connected to the intended concept (Taherdoost, 2016). Despite criticisms of face validity as being 'trivial' and 'cosmetic,' its incorporation into the citizen participation process is deemed beneficial by researchers (Bannigan & Watson, 2009; Patton, 2002). For policymakers to regard PVE as a legitimate source in decision-making processes, it is essential that the participants perceive its authenticity and utility.

Despite the importance of face validity, studies on PVE in this regard remain underexplored. Understanding why participants evaluate the face validity of PVE is essential, particularly regarding inclusiveness and sustainability in transportation planning. Questions arise about whether certain groups (in)consistently perceive PVE as effective and if the online format of PVE equally engages all citizens. Addressing these questions requires segmentation of respondents based on demographics, behaviors, or attitudes to tailor engagement strategies and improve the face validity of PVE. Analyzing participation results at a granular level enables organizations to identify trends, preferences, and concerns specific to each segment, thereby enhancing decision-making (Barnett & Mahony, 2011). By identifying groups that consistently evaluate PVE negatively and the reasons for their negative evaluations, future PVE designs can be developed with tailored treatments to increase its face validity, thereby increasing its effectiveness as a public participation tool.

1.2 Knowledge Gap and Main Research Questions

Transportation plays a pivotal role in society, offering both benefits and drawbacks to the environment and communities. To manage these effects, transportation planning aims to strike a balance between the positive and negative impacts of transportation activities. However, contemporary transportation planning faces evolving challenges beyond traditional concerns like reducing travel time and traffic accidents. There is a growing emphasis on inclusiveness and forward-looking goals, which traditional tools like SCBA alone may struggle to address effectively.

In response to these challenges, Participatory Value Evaluation (PVE) emerges as a promising approach. PVE offers a way to incorporate considerations related to future transportation goals that

traditional methods may overlook. Moreover, by involving citizen participation, PVE can enhance decision-making for transport planning practitioners and governments. However, the effectiveness of PVE hinges on its validity, particularly face validity.

While some studies have explored the face validity of PVE, there is a notable gap in understanding how groups of citizens perceive these evaluations. Are there specific characteristics or groups of citizens that consistently assess the face validity of PVE in particular ways? If a significant portion of the population evaluates PVE negatively, its legitimacy as a decision-making tool for transport planning practitioners could be undermined. Thus, segmentation becomes crucial in analyzing the face validity evaluation of PVE, as it can inform tailored strategies for improving its validity in future designs.

The lack of research on the segmentation of PVE's face validity evaluation in transportation planning leads to the following main research questions:

"To what extent are there distinct segments in the population who (in)consistently evaluate the face validity of Participatory Value Evaluation (PVE) in transportation planning?"

In addition, supplementary questions emerged to complement the main research question:

- What are the factors contributing to the negative assessment of PVE by these specific segments?
- What strategies can be implemented to improve the evaluation of PVE's face validity for segments that consistently express negative views towards it?

1.3 Relevance of the Research

This research contributes significantly to the scientific discourse by delving into the validity of the PVE method, an area that has received limited attention thus far. By specifically investigating the face validity of PVE, this study sheds light on citizens' expectations and evaluations regarding public participation within the PVE framework. This insight allows for future PVE to better accommodate citizens, enabling them to provide more informed advise to policymakers and improving the effectiveness of their participation. Thus, the primary objective of this research is to identify the characteristics of citizens who (in)consistently evaluate the face validity of PVE negatively and propose recommendations for accommodating these groups in future PVE initiatives to improve their face validity evaluation.

The scientific significance of this study lies in the validity of a novel research tool and its contribution to the field of public participation in transportation planning. By validating the PVE method, this research enhances its credibility and utility as a decision-making tool, ultimately enriching the practice of transportation planning. In addition, this study aligns with TNO's Safety Urban Mobility and Safety (SUMS) research, which aims to produce inclusive and forward-looking transportation systems. This approach is in accordance with welfare beyond GDP perspective, also known as *"brede welvaart"* in the Dutch context.

Additionally, this research holds relevance for the Transport, Infrastructure, and Logistics (TIL) master's program, particularly within the Transport Governance specialization. The focus of Transport Governance in TIL centers on the transportation planning process, which culminates in the formulation of transport policy plans and proposals for transport and logistics projects, along with their decision-making. This study directly addresses the validity of PVE as a promising new method in transportation planning, offering actionable insights and recommendations for future PVE designs. Consequently, it contributes to the efficacy of PVE as a tool to assist transport planning practitioners and governments in decision-making processes.

1.4 Research Approach and Sub-questions

In order to address the objective mentioned previously, several steps will be taken in this research. First, the face validity and overall ratings of the PVE consultations quantitative analysis will be conducted. The quantitative analysis of face validity statements are based on under development instrument that has been initiated by Tuit (2022)'s thesis and de Ruijter (2023)'s study. Following that, qualitative evaluation is conducted to help explain the quantitative data in terms of why participants evaluated the overall ratings as they did and what they did not like from the consultation. Finally, the findings

from these previous steps will be discussed with PVE experts to draw conclusion and discuss practical solution for enhancing the future PVE consultation.

To answer main research questions, some sub-questions need to be answered as followed.

1. **To what extent are there distinct segments in the population who (in)consistently evaluate the face validity of Participatory Value Evaluation (PVE) in transportation planning?**

   1.1. *What is the current state of knowledge, in literature, regarding face validity in public participation approaches, particularly within the realm of transportation planning?*

   This study revolves around the face validity of PVE, as one public participation approach in transportation planning. Therefore, it is important to investigate what face validity exactly contains and how face validity has been measured or questioned in other studies, especially in the transportation field. To answer this sub question, a comprehensive literature review is performed.

   1.2. *What are the characteristics of the distinct segments in PVE consultations, and how do these characteristics relate to their face validity evaluations?*

   To determine if there are specific groups in society that possess specific characteristics and attitudes towards the face validity of PVE, it is imperative to understand how respondents assess the face validity of PVE and what their individual characteristics are. The dataset obtained from the consultations of Lelylijn, Oude Lijn, and Mobility Vision were utilized to address this inquiry. Initially, descriptive statistics analysis were conducted to examine the features of the sample from all three consultations. Additionally, descriptive statistics were conducted to analyze the participants' assessment of each face validity statements in each consultations. Subsequently, crosstabulation analysis were used to examine the correlation between the respondents' attributes and each face validity statement. Following that, Latent Class Cluster Analysis were undertaken to see if there are certain groups within the population, depending on the features of the respondents and their appraisal of face validity. Ultimately, Multinomial Logistic Regressions were conducted to determine the specific features of individuals that impact their assessment of the face validity of the three PVE consultations.

   1.3. *How consistent are these face validity evaluations across different PVE experiments in transportation planning?*

   To ensure the generalizability of the findings in sub-question 2 regarding the characteristics of respondents and the emergence of distinctive segments within the population, the results of the mentioned quantitative analysis of all the 3 PVE consultations were compared. Based on this comparison, a conclusion was drawn regarding the varying assessment of the face validity of PVE by specific groups.

2. **What are the factors contributing to the negative assessment of PVE by these specific segments?**
   It is crucial to recognize that there is a particular group that constantly evaluates PVE adversely, and it is important to understand the reasons behind their negative evaluation. A content analysis was conducted to determine the underlying reasons behind the respondents' evaluation of the face validity of PVE Lelylijn. This research focused on the respondents' answers regarding their dislikes about the consultation process.

3. **What strategies can be implemented to improve the evaluation of PVE's face validity for segments that consistently express negative views towards it?**
   To answer the above research question, findings from the previous research questions are shared with PVE experts, in a focused group meeting interviews. Conclusions are drawn and practical solutions for enhancing future PVE consultations in transportation planning are discussed.

An overview of all methods that are used is presented in Figure 1.2.

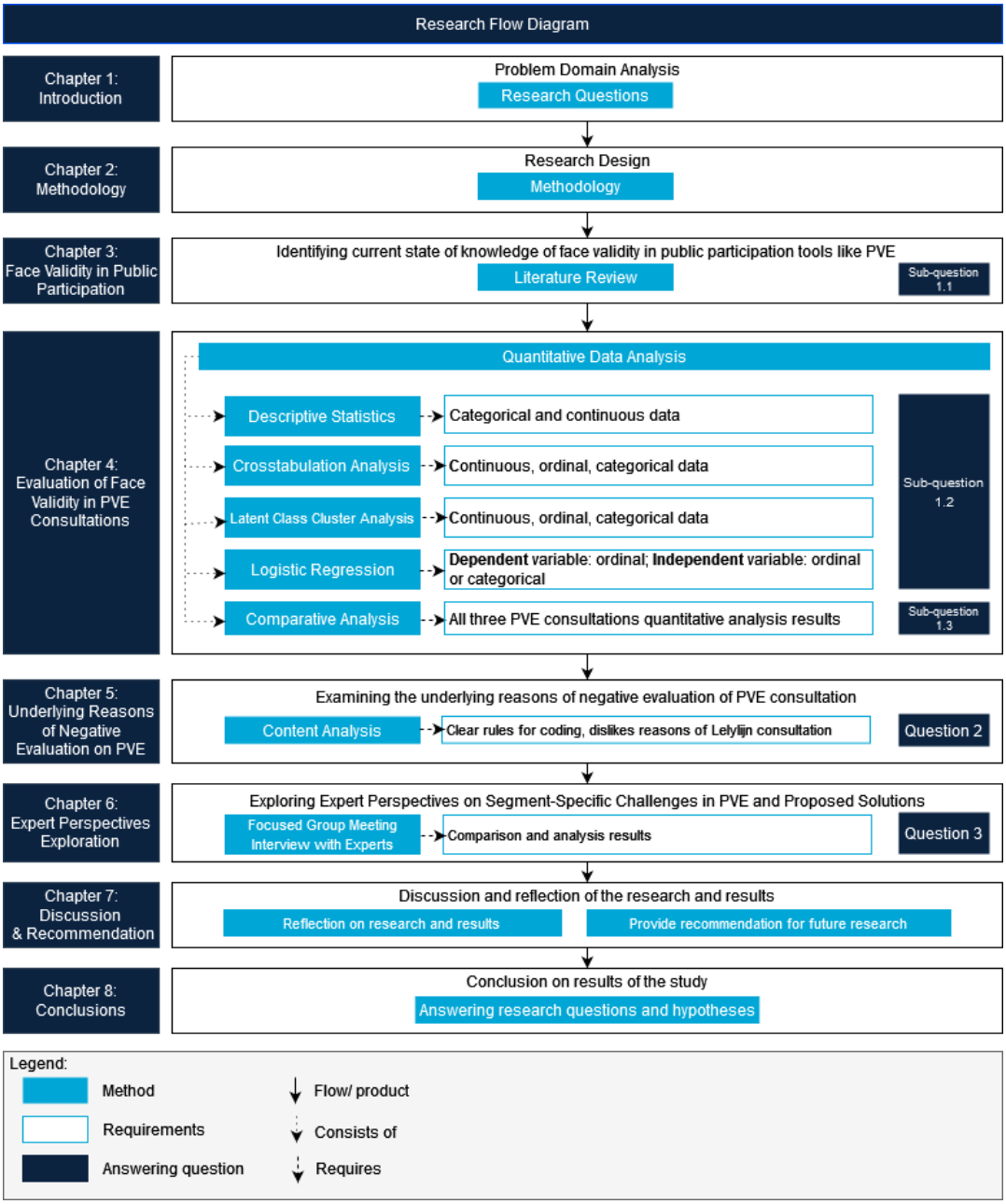


Figure 1.2: Thesis Research Flow

1.5 Summary and Structure of the Report

The structure of this report is as follows. The next chapter, Chapter 2, elaborates on the research methods applied in this study. In Chapter 3, Sub-question 1.1 is elaborated as the results from the literature review. Following that, Chapter 4 provides the results of the quantitative analysis of all the three PVE consultations and discusses if there are (in)consistent trends across datasets. Next, Chapter 5 presents the results of the content analysis of Lelylijn consultation. Chapter 6 discusses the findings from the focus group meeting with experts. Chapter 7 provides the discussion and recommendations. Finally, Chapter 8 elaborates on the conclusion for this study.

Key Takeaways

- Transportation is essential for societal functionality, facilitating movement while also presenting environmental and safety challenges. Effective transportation planning aims to maximize benefits and minimize negative impacts.
- There is a shift from traditional one-way decision-making in transportation planning to more inclusive and deliberative processes, such as Participatory Value Evaluation (PVE), to better incorporate public input and societal values.
- PVE is an innovative online tool allowing citizens to engage in the policymaking process by choosing between multiple policy options, reflecting both individual and societal preferences.
- PVE aligns with broad welfare perspective
- Assessing the face validity of PVE is important to ensure its credibility and effectiveness as a tool for public participation. Understanding how different groups perceive PVE (in)consistently is crucial for improving its design and implementation.
- Key **research questions** aimed at understanding how different segments of the population evaluate PVE's face validity and how to enhance its effectiveness as a participatory tool.

2 Methodology

This chapter describes the case studies and methods that were performed to answer the research questions. Table 2.1 presents an overview of the sub-questions and methods used.

Table 2.1: Research Questions and Methodologies used to Answer Them

No		Research Question	Method
1	To what extent are there distinct segments in the population who (in)consistently evaluate the face validity of Participatory Value Evaluation (PVE) in transportation planning?		
	1.1	What is the current state of knowledge in literature, regarding face validity in public participation approaches, particularly within the realm of transportation planning?	Literature Review
	1.2	What are the characteristics of the distinct segments in PVE consultations, and how do these characteristics relate to their face validity evaluations?	Descriptive Statistics Crosstabulation Analysis Latent Class Cluster Analysis Logistics Regression
	1.3	How consistent are these face validity evaluations across different PVE experiments in transportation planning?	Comparative Analysis
2	What are the factors contributing to the negative assessment of PVE by these specific segments?		Content Analysis
3	What strategies can be implemented to improve the evaluation of PVE's face validity for segments that consistently express negative views towards it?		Experts Group Meeting

This study utilized three different PVE project consultation datasets, as outlined in subsection 2.1. The selection of these consultations was based on several factors. Firstly, these consultations provide extensive, recently collected, and compiled datasets. Furthermore, the three consultations aim to address the issues of transportation and accessibility, which are currently significant concerns for the Netherlands government. Additionally, the descriptive results in section 4.2 show comparable outcomes across all three datasets, indicating a high level of reliability. This consistency enhances the explanatory power of the comparative analysis, making the comparison of the datasets more valid and meaningful. Therefore, these characteristics make these consultations suitable for this research and align with the objectives of the Master's TIL Transport Governance specialization and TNO's Sustainable Urban Mobility & Safety (SUMS) focus.

The PVE consultations incorporated questionnaires designed to evaluate the effectiveness of PVE as a methodology. The questionnaires, which are anonymous, include demographic data and participant characteristics, as well as a set of questions about the overall rating and face validity statements of the PVE consultation.

2.1 Case Studies and Data Collection

The data used in this study were collected by different researchers and were not part of this study's original data collection process. Instead, these datasets were obtained from recently conducted Participatory Value Evaluation (PVE) consultations in the Netherlands. These PVE consultations were intentionally designed with slight variations to cater to the distinct objectives of each consultation. Nevertheless, the consultations consist primarily of several sections that are consistent across all consultations.

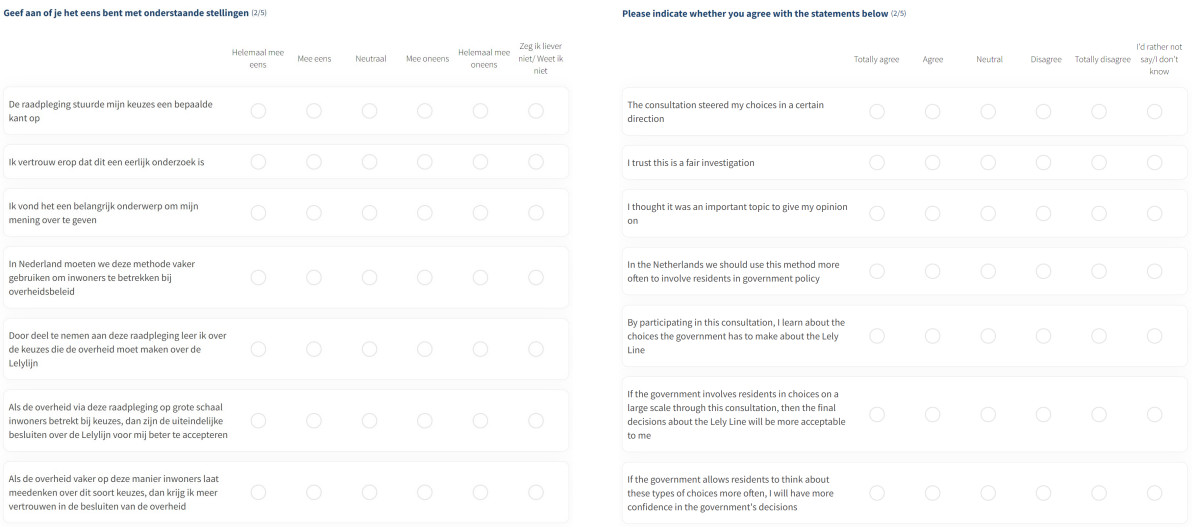
These sections include a segment dedicated to policies selection (i.e., point allocations or slider choice task) and the motivations, a section focused on gathering information about the sociodemographic characteristics of the respondents, a section dedicated to the respondents' review of the consultation, and an additional section that is customized for each specific consultation. Although there were some questions regarding the respondent's characteristics, the consultations were entirely anonymous, and the responses could not be used to identify individuals. This research specifically examines the face validity evaluation of PVE. As a result, the study solely employed respondents' sociodemographic data and their reviews of the PVE consultation.

The case studies employed in this research are further explained as follows.

2.1.1 Lelylijn Project

Lelylijn project is a proposed railway line in the Netherlands designed to link the Northern Netherlands region with the rest of the country. The initiative aims to enhance transportation infrastructure in the Northern Netherlands to facilitate residents' access to crucial amenities and to increase the region's attractiveness to young individuals and improve its economic opportunities.

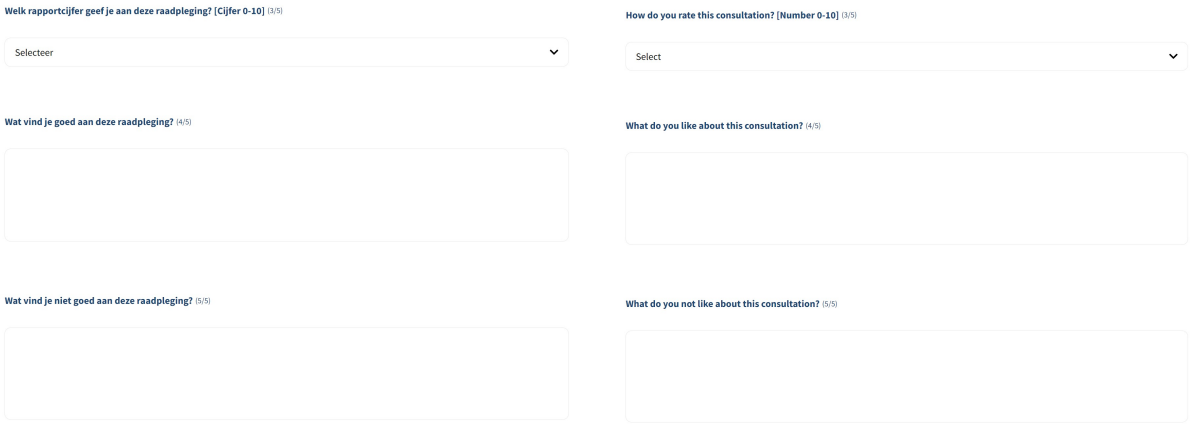
The PVE consultation data for the Lelylijn project was acquired in 2023 and the PVE project analysis has been completed. Nonetheless, the data on face validity and participants' consultation experience has the potential to be further researched to enhance the usefulness of PVE as a tool for transport decision-making. The data consists of participants' socio-demographic information, decision tasks throughout the consultation, and participants' ratings on the consultation procedure and the experiment's face validity. Figure 2.1 and 2.2 show the list of statements and overall rating questions participants see in the PVE Lelylijn consultation.



(a) Statements in Dutch

(b) Statements in English

Figure 2.1: Statements for Face Validity in Lelylijn Project PVE Consultation. Aligned with Face Validity instrument in Tuit (2022) thesis.



(a) Questions about rating in Dutch

(b) Questions about rating in English

Figure 2.2: Questions asked in Lelylijn PVE Consultation for participant's overall rating and open question to be analysed in content analysis 2.2.8.

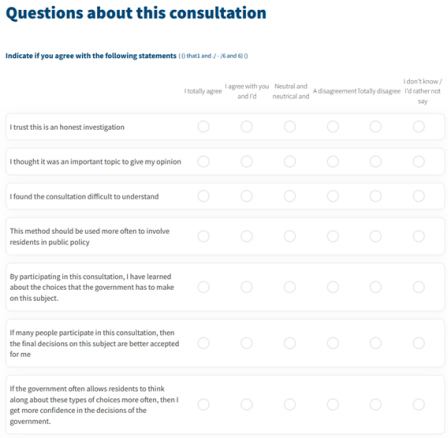
2.1.2 Oude Lijn Project

The Oude Lijn is a historic railway line that has connected the cities of Leiden and Dordrecht since 1847, running through Leiden, The Hague, Delft, Schiedam, Rotterdam, and Dordrecht. This railway is crucial for the Southern Randstad region, which is expected to experience substantial economic, population, and infrastructure growth in the coming years. With plans to add 170,000 housing units and 85,000 jobs in the cities along the Oude Lijn, there is a critical need to enhance accessibility while preserving the region's green spaces (Ministerie van Infrastructuur en Waterstaat, 2024).

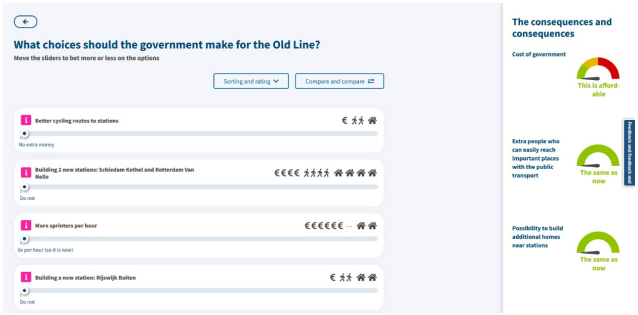
To facilitate this regional development, the Ministry of Infrastructure and Water Management, the Province of South Holland, the Rotterdam-The Hague Metropolitan Region, ProRail, and the municipalities of Leiden, The Hague, Schiedam, and Dordrecht have embarked on a MIRT exploration. MIRT stands for Multi-Year Program for Infrastructure, Space, and Transport. This exploration aims to collaboratively develop smart, sustainable, and climate-resilient solutions to the challenges along the Oude Lijn.

The MIRT exploration comprises six subprojects, focusing on key nodes: Leiden Central, The Hague Laan van NOI, Schiedam Centrum, and Dordrecht. Each project ensures that the spaces around these stations can accommodate future urban growth. ProRail's subproject focuses on additional (train) sprinters and new stations, while another joint project addresses the coordination between spatial and mobility developments.

The Oude Lijn project's PVE consultation was chosen as one of the datasets for this thesis due to its significance as a transportation project that incorporates public participation. While the Lelylijn project has not yet been implemented, the Oude Lijn railway line is already in place. Hence, utilizing the Oude Lijn dataset is beneficial for assessing the respondents' evaluation of the PVE since it can add additional perspective of PVE evaluations on project on a project that has been already existing.



(a) Face Validity Statements in Oude Lijn Consultation

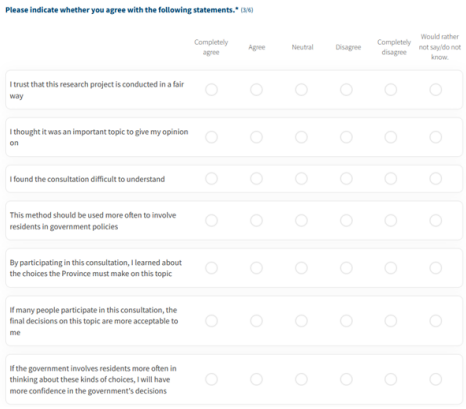


(b) Visualization from PVE Oude Lijn

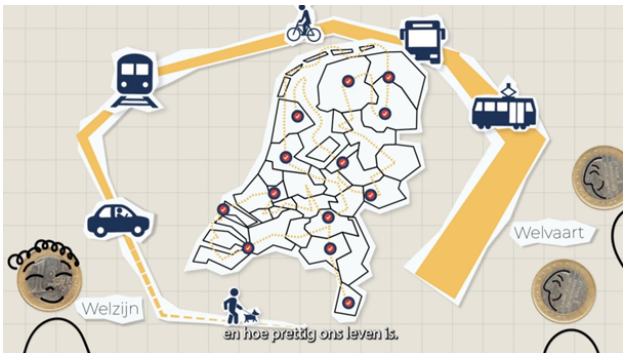
Figure 2.3: Oude Lijn Consultation

2.1.3 Mobility Vision Project

The Mobility Vision project is crucial for understanding the forthcoming developments in transportation in the Netherlands. The government's comprehensive plan aims to ensure effective transportation for all individuals, irrespective of their chosen mode of travel. The mobility vision PVE consultation aimed to gather public input on the government's strategy for decision-making on mobility, taking into account the constraints of limited resources and space. This strategy prioritizes the significance of harmonizing different forms of transportation to maximize their advantages and guarantee smooth and uninterrupted communication (van Infrastructuur en Waterstaat, 2020).



(a) Face Validity Statements in English



(b) Visualization from PVE Mobility Vision

Figure 2.4: Mobility Vision consultation

The dataset was selected for study in the thesis because to its representation of a future-oriented approach to transportation planning in the Netherlands. The study seeks to comprehend how various groups of people perceive and assess the face validity of PVE within the framework of a comprehensive, future-oriented transportation planning endeavor through the analysis of this dataset. This understanding is crucial for improving the credibility and efficiency of public engagement methods in influencing future transportation policies.

2.2 Methods

The following sections explain the methodologies utilized in this thesis. The numerical data obtained from the PVE consultation of the case studies were utilized to perform statistical analysis in order to resolve the research questions. The PVE consultations provided extensive data from panel and open

consultations, meeting the minimum sample size requirements specified by Knofczynski and Mundfrom (2007). This study utilized a multi-method approach, which included doing a literature review, statistical analyses, content analyses and group discussion with experts, in order to thoroughly evaluate the face validity of Participatory Value Evaluation (PVE). These methodologies were chosen to address both the quantitative and qualitative parts of the research. The conceptual framework of research methodology of this study is presented in Figure 2.5. Statistical analysis was performed using SPSS and LatentGold 6.0. Additionally, other methodologies were employed, all of which are described below.

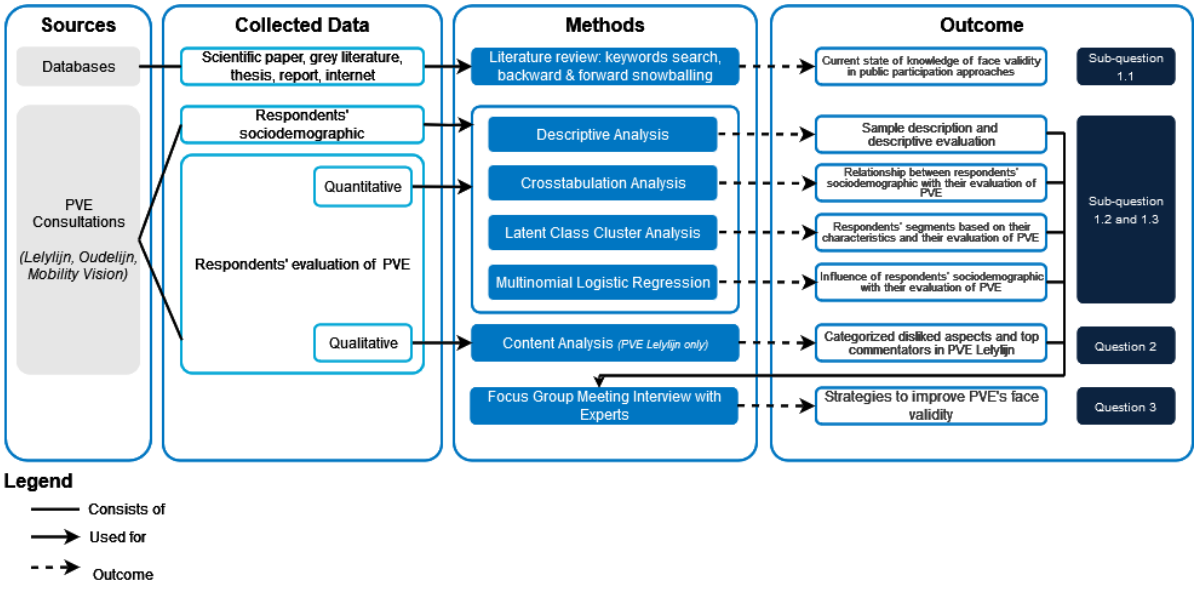


Figure 2.5: Conceptual Framework of Research Methodology

2.2.1 Literature Review

A literature review is an essential technique in academic research that involves identification, assessment, and integration of existing research on a certain subject. This approach enables researchers to gain a thorough understanding of the existing knowledge, identify gaps in the literature, and situate their own research within the wider academic framework (Snyder, 2019). By analyzing and summarizing a wide range of sources, literature reviews provide a critical overview that informs and guides subsequent research efforts.

In this study, the databases Scopus and Google Scholar were utilized as sources for references. The search keywords utilized were: citizen participation, participatory value evaluation, PVE, validity of research instrument, and face validity. Additionally, the master theses of Tuit (2022) and Golan (2023) were utilized for backward snowballing to provide an overview of the current condition of PVE. Forward snowballing were also undertaken from Taherdoost (2016) and Gaber and Gaber (2010) to determine the current research on validity of research instrument. Articles were chosen based on their relevance to public participation in decision-making, particularly the PVE approach in transportation, and the validity of preference elicitation experiments. This comprehensive analysis of scientific and grey literature aims to identify areas of consensus and gaps in existing knowledge are provided in Table 3.1 in Chapter 3.

2.2.2 Data Cleaning

The data obtained from the three PVE consultations consist of a lot of information that were not all needed in this study. Hence, data cleaning, as depicted in Figure 2.6, were performed to only have the respondents' sociodemographic data, their evaluation of face validity, and their ratings for overall PVE consultation. For Lelylijn dataset only, the respondents' response on what they did not like about the consultation, was still retained, for the content analysis (see section 5). The data were also filtered to exclude the "I don't know/ I rather not say", for consistency purposes. Moreover, the data utilized in this study were only one with complete answers (i.e., if at least one out of seven face validity statements was not answered, the whole response was deleted from the dataset). Therefore, the total data and

and descriptive results of this study might be different than the actual report of the PVE consultations.

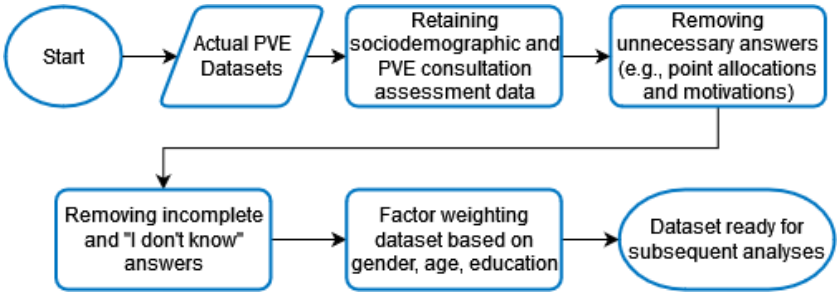


Figure 2.6: Data Cleaning and Processing

Following data cleaning, a descriptive analysis was performed on key respondent characteristics (age, gender, and education level) to assess whether the sample proportions were comparable to those of the Dutch population. This initial analysis revealed that the demographic profile of respondents in all three consultations differed from the general population in terms of gender, age, and education. Preliminary statistical analyses, which did not account for these demographic differences, yielded less significant and meaningful conclusions. Consequently, factor weighting was applied to the datasets to adjust for these discrepancies in subsequent analyses.

The factor weighting process started with calculating the Chi-square statistic using formula (B.1) and determining the degrees of freedom (dof) with formula (B.2). The critical Chi-square value for the given dof was then identified using a Chi-square distribution table. Subsequently, a weight factor for each attribute level was calculated using formula (B.3).

- Calculate the chi-square statistic using the formula:

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} \tag{2.1}$$

  where O_i is the observed frequency and E_i is the expected frequency for each cell.

- Determine the degrees of freedom for the test, which is calculated as:

$$(\text{rows} - 1) \times (\text{columns} - 1) \tag{2.2}$$

- Look up the critical chi-square value in a chi-square distribution table or use statistical software to find the p-value associated with the chi-square statistic.
- Calculate a weight factor for each attribute level using the formula:

$$\text{Weight Factor} = \frac{E_i}{O_i} \tag{2.3}$$

Factor weighting was performed using Microsoft Excel and SPSS. After the first iteration, another Chi-square test was conducted to assess whether the adjusted sample proportions matched those of the Dutch population. This process was repeated until the sample proportions were statistically comparable to the population proportions.

2.2.3 Descriptive Analysis

Descriptive analysis were utilized to characterize the population sample and provide a detailed description of the collected data (Delaney, 2010). This involved summarizing the quantitative data from the PVE consultation to understand the distribution of responses. General demographic characteristics such as age, gender, and education level, as well as case-specific characteristics like participants' proximity to the study site, were analyzed. This analysis helped provide insight into how each face validity statement and the overall consultation rating were scored in each consultation. Furthermore, this research used to determine if the composition of the sample is comparable to that of the population, which influenced the decision to employ factor-weighting for the datasets.

2.2.4 Crosstabulation Analysis

Crosstabulation analysis, commonly referred to as contingency tables, is a statistical technique employed to investigate the relationship between categorical variables through the grouping and comparison of data. This methodology enables researchers to analyze the relationships between various variables and detect patterns or interdependencies within the dataset. Crosstabulation is a method of putting data into a table style to provide a concise and comprehensive understanding of how factors interact with each other (Mohd Pakri et al., 2024).

In this study, crosstabulation analyses were used to investigate the associations between the characteristics of the respondents (e.g., age, gender) with their assessments of PVE consultations (i.e., ratings of face validity statements and overall PVE consultation).

To determine the significance and strength of the association between the respondent's characteristics and their evaluation on the PVE consultation, two statistical factors were employed. The Gamma coefficient was employed to measure the strength and direction of these associations, with values ranging from -1 to 1. A positive Gamma value indicates a direct relationship, while a significance (Sig.) value below 0.05 was used to determine statistical significance, indicating that the observed relationships are unlikely to be due to chance. (IBM, 2024).

2.2.5 Latent Class Cluster Analysis (LCCA)

A Latent Class Cluster Analysis (LCCA) was performed to identify distinct groups of participants who exhibit either positive or negative evaluation on certain statements related to face validity and consultation assessments in each dataset. Positive evaluation is indicated by agreement with positively formulated face validity statements and high consultation ratings, while negative evaluation is indicated by disagreement with positively formulated statements or agreement with negatively formulated statements and poor consultation ratings. The data needed for the LCCA must fall into one of the following categories: ordinal, continuous, or count. Furthermore, it is assumed that the indicators in an LCCA are not influenced by each other (local independence assumption) (Tuit, 2022).

The analysis used categorical face validity responses (”agree,” ”neutral,” ”disagree”) as indicators, alongside overall consultation ratings. Sociodemographic characteristics were included as covariates. Models with 1 to 10 clusters were tested, with the optimal model selected based on several criteria. The Log-Likelihood (LL) assesses the model's fit to the data, with higher values indicating a better fit. Information criteria such as the Bayesian Information Criterion (BIC), the Akaike Information Criterion (AIC), and the AIC with a penalty for complexity (AIC3) were also utilized, where lower values suggest more parsimonious models. The Maximum Bivariate Residual (Max. BVR) evaluates the largest discrepancy in the pairwise associations, with lower values indicating better model fit. The classification error rate measures the accuracy of the cluster assignments, with lower values being preferable. Lastly, the Entropy R^2 assesses the precision of the classification, with higher values denoting greater separation between clusters and clearer cluster distinctions (Kroesen, 2024; Vermunt & Magidson, 2002). Besides that, the interpretability and meaningfulness of the clusters results were also considered in choosing the optimal model.

2.2.6 Logistic Regression

Initially, Ordinal Logistic Regression was planned to explore the potential influence of respondents' characteristics on their evaluations of face validity within different categories. This analytical approach was chosen considering the ordinal nature of the dependent variable, represented by Likert scale categories. However, meeting the parallel regression assumption is essential for this method, as it indicates a linear relationship between each independent variable and the logit of the dependent variable.

Parallel line tests were conducted to evaluate the proportional odds assumption before performing Ordinal Logistic Regression. The objective of this test is to determine whether the assumption of proportional odds holds true in an ordinal logistic regression model. Specifically, in this context, the outcome groups (i.e., ”agree,” ”neutral,” and ”disagree”) were examined. This test assesses if the relationship between each pair of these outcome groups is statistically the same, indicating that the effect of the predictor variables is consistent across different threshold levels of the ordinal outcome group.

The significance of the chi-square statistic in the test of parallel lines is used to check this assumption. A significant p-value (less than 0.05) suggests that the assumption of proportional odds (parallel lines)

is violated. In other words, the relationship between the predictors and the outcome variable differs across the response categories of ”agree,” ”neutral,” and ”disagree”. This violation indicates that the proportional odds assumption may not be appropriate for the data, and hence the use of Ordinal Logistic Regression might not be suitable (Liang et al., 2020).

As a result, Multinomial Logistic Regression was employed to explore the potential influence of respondents' characteristics on their evaluations of face validity within different categories. This method was chosen due to its ability to handle non-linear relationships without imposing the strict parallel regression assumption. In this study, the dependent variable was the ordinal face validity ratings, represented by Likert scale categories (agree, neutral, disagree). The independent variables included demographic characteristics such as age, gender, and education level, as well as case-specific characteristics like participants' proximity to the study site and previous experience with PVE consultations.

2.2.7 Comparative Analysis

A comparative analysis was performed to investigate the consistency of respondents' evaluation, specifically addressing the third sub-question. The objective was to enhance the understanding of how segment groups consistently evaluated the face validity and overall rating in different PVE consultations in transportation field. This was accomplished by doing a comparative analysis of clusters and respondents' characteristics across various PVE consultations in order to see if particular segment groups consistently exhibited similar evaluations of face validity. In addition, the findings of this study are also compared to the results of the prior face validity study and the hypotheses made in Chapter 3.3.

2.2.8 Content Analysis

The study utilized content analysis, which involved a thorough examination of textual data using both qualitative and quantitative methods, as described by Aacharya (2022). This adaptable method was crucial in assessing the written justification behind participants' assessment for the overall PVE consultation, hence addressing the research question 2. This involved investigating the components of the PVE consultation that participants found unappealing. The process involved categorizing text into manageable sections for analysis, thus complementing the assessment of overall ratings. Although content analysis is a time-consuming process and has the possibility of mistakes, it provides vital historical and cultural insights by combining qualitative and quantitative analysis in a nuanced way (Columbia University, 2023).

This study employed content analysis to examine participants' responses about concerns regarding the consultation, as depicted in Figures 2.2a and 2.2b. Given the complexities and resource demands of content analysis, this method was selectively applied to the Lelylijn negative feedback dataset, which was both the first consultation conducted and the earliest available, providing timely access to the richest and most detailed qualitative feedback.

2.2.9 Focus Group Meeting Interview with Experts

To address the third research question, a focus group meeting interview was conducted with PVE experts. Expert interviews are crucial for obtaining up-to-date insights and enhancing understanding of the research topic (Morris et al., 2018). The insights and patterns gathered from earlier stages were deliberated with PVE experts to explore potential practical solutions aimed at improving public perceptions of PVE consultations. Engaging PVE experts is advantageous due to their familiarity with study design, data interpretation, and extensive experience in PVE studies. Their input is crucial in identifying actionable steps for enhancing future PVE efforts based on the findings of this study.

2.3 Summary

Datasets and Methods

- The research utilizes datasets from three recent PVE consultations in the Netherlands: Lelylijn, Oude Lijn, and Mobility Vision projects, each providing comprehensive data on participant demographics, decision tasks, and face validity evaluations.
- Data from the PVE consultations were cleaned and filtered to focus on respondents' sociodemographic data and their reviews of the PVE consultation. **Factor weighting** of the dataset was applied to adjust for demographic discrepancies.
- **Descriptive and Crosstabulation Analysis** were used to characterize the population sample, understand response distributions, and explore relationships between respondent characteristics and PVE evaluations.
- **Latent Class Cluster Analysis (LCCA)** was performed to identify distinct groups of participants based on their characteristics and PVE evaluations
- **Multinomial Logistic Regression** was used to explore the influence of respondent characteristics on face validity evaluations
- **Comparative Analysis** on the respondents' characteristics and evaluations across different PVE consultations were conducted to identify (in)consistent patterns.
- **Content Analysis** was applied to the Lelylijn dataset, for understanding underlying concerns regarding the PVE consultation.
- **Focus group meetings with PVE experts** were conducted to discuss findings and develop practical solutions for improving future PVE consultations.

3 Face Validity and Evaluation of PVE

This chapter provides the results of a literature review that examines the current knowledge on the face validity of public participation approaches, with a specific focus on transportation planning. Therefore, it provides a response to sub-question 1.1, which is *"What is the current state of knowledge in literature, regarding face validity in public participation approaches, particularly within the realm of transportation planning?"*. First, it provides a detailed explanation of the concept of face validity. Subsequently, the chapter discusses the current understanding of face validity in public participation. Additionally, it explores hypotheses related to PVE evaluation.

3.1 Face Validity

The evaluation of public participation's success hinges on the crucial aspect of assessing its effectiveness, as emphasized by Rosener (1978). Effectiveness, as defined by Rowe and Frewer (2004), is the measure to which an instrument or process adequately gauges the goal concept, introducing the concept of validity. Validity, the extent to which an instrument measures its intended construct, is a nuanced concept relative not only to the context of a specific research project but also within the types of data generated and the methods employed (Boxebeld et al., 2024; Gaber & Gaber, 2010; Tuit, 2022).

Despite its significance, validity often goes unaddressed for commonly used instruments due to the extensive work required (Bannigan & Watson, 2009). Many questionnaire studies proceed without validity checks due to various constraints, such as a lack of knowledge, time constraints, complex research schemes, exceptional questionnaires, and limited resources (Patel & Desai, 2020). However, if an instrument, like a PVE consultation, lacks validity, there's a risk of obtaining misleading results which potentially leading to misguided decisions (Anastasi & Urbina, 2007).

Various approaches are recommended to test validity, with the suggestion that multiple validation procedures should be employed rather than relying on a single method (Bannigan & Watson, 2009). A quickest one to test the validity of an instrument is face validity which assess whether on the face of the instrument, the research makes sense. Taherdoost (2016) defines face validity as the extent to which a measure seems to be connected to a particular concept, as perceived by individuals who are not experts in the field, including test takers representatives of the legal system. However, critiques persist about face validity, with some arguing its observations are 'trivial' and 'cosmetic' since they lack empirically verifiable testing procedures (Gaber & Gaber, 2010). Moreover, some researchers dispute face validity, noting that respondents may perceive an instrument as valid even when it fails to measure what it is supposed to (Oluwatayo, 2012).

Despite these critiques, introducing face validity into the citizen participation process is deemed beneficial by researchers. Patton (2002) suggests that incorporating face validity can enhance planners' understanding of the 'interaction and mutuality' of proposed plans on communities directly impacted by them (Gaber & Gaber, 2010). Bannigan and Watson (2009) further highlights the importance of assessing face validity. When an instrument is considered face-valid by respondents and stakeholders, it ensures the acceptance and usefulness of the instrument. In the context of the PVE method, ensuring its perceived authenticity and utility is crucial for policymakers to consider it a legitimate source in decision-making processes.

3.2 Current State of Knowledge on Face Validity in Public Participation Instrument

The notion of validity is essential in the construction and evaluation of a measurement instrument. According to Ghauri et al. (2020), validity is the extent to which the data gathered by an instrument precisely addresses the particular subject under examination. Taherdoost (2016) differentiated between several types of validity, including face validity, content validity, construct validity, and criterion validity.

Ensuring the credibility of public involvement methods, such as Participatory Value Evaluation (PVE), is crucial since any perceived lack of validity can greatly affect the method's credibility, appeal, and acceptability (de Ruijter, 2022).

Face validity, a crucial aspect of validity, refers to the degree to which a test seems to measure what it claims to measure (Tuit, 2022). This form of validity is crucial for guaranteeing that the tool is regarded as relevant, easy to understand, and acceptable by its target audience. The importance of face validity lies in its impact on the credibility and acceptance of the instrument by participants. When participants view an instrument as having face validity, they are more inclined to genuinely engage with it and deliver accurate responses. This emphasizes the significance of implementing face validity evaluations in diverse domains.

Application of Face Validity in Different Fields

In healthcare, face validity assessments have been used to evaluate the clarity and comprehensibility of questionnaire items. For instance, the MUAPHQ C-19 questionnaire, developed to assess understanding, attitudes, practices, and health literacy regarding COVID-19 among Malaysian adults, was found to be relevant and understandable by participants. This underscores the importance of cultural and linguistic considerations in public health tools (Dalawi et al., 2023; Reffien et al., 2022). Similarly, the Evaluation Tool of Health Information for Consumers (ETHIC), designed to assess the quality of health information materials in Italian, aimed to measure the tool's usability and comprehensibility among potential users (Hermans et al., 2016; Morton et al., 2023). Another example is the face validity assessment of a physical activity questionnaire for Spanish-speaking women in California, which emphasized cultural appropriateness and understandability (Banna et al., 2023).

Face validity is employed in the field of environmental and agricultural economics to evaluate the accuracy of willingness to pay (WTP) estimations derived from choice experiments. Experts evaluate the rationality of WTP estimates, detecting instances of 'overshooting' where the projections beyond reasonable thresholds. The purpose of this assessment is to prevent mistakes that could lead to inaccurate policy decisions or ineffective solutions for environmental management (Glenk et al., 2024). This exemplifies the pragmatic application of face validity in assessing the reliability of research instruments and ensuring that estimated values correspond to rational expectations in decision-making.

In financial risk tolerance assessments, face validity is critical for ensuring the effectiveness of financial measurement tools. A study in Malaysia evaluated the financial risk tolerance of investors with different levels of investment expertise, involving lay experts to verify the clarity, comprehensibility, and relevance of the questionnaire items. This approach improves the quality, accuracy, and user-friendliness of the instrument, leading to more dependable and legitimate measurement results (Hadi et al., 2023).

Face Validity in Public Participation Methods

Public participation initiatives generate various tangible benefits for participants and society, including the acquisition of dignity, self-esteem, and the opportunity to contribute meaningfully to decision-making processes (Burton, 2009). However, to realize these benefits effectively, it is essential that public participation processes and instruments exhibit face validity. Participants must view the measures and procedures applied in public participation as clear, relevant, and aligned with their experiences. Ensuring face validity in public participation is crucial for fostering trust, engagement, and meaningful involvement, ultimately enhancing the quality and impact of the evaluation outcomes.

Participatory Value Evaluation (PVE) is a new approach used to gather public preferences in the transportation sector. Tuit (2022) initiated a face validity measurement framework for PVE, which consists of a series of statements designed to assess the clarity, relevance, and comprehensibility of the approach. The purpose of the face validity assessment is to evaluate participants' judgments regarding the instrument's suitability and efficacy in accurately collecting their opinions on policy options. Golan (2023) conducted a comparative analysis of the face validity evaluation of public value elicitation (PVE) for public transport policy preferences in Tel Aviv, Israel, and previous PVEs conducted in the Netherlands. The findings indicated that both studies reported similar ratings of lower completeness aspect relative to other features. de Ruijter (2023) provides a concise summary of the most recent progress

regarding assessing the face validity of PVE. She presents a face validity assessment tool that has nine distinct dimensions.

In conclusion, face validity is a critical component in the development and implementation of public participation instruments. Ensuring that these tools are perceived as relevant, clear, and acceptable by the target audience enhances their effectiveness and acceptance. The reviewed studies highlight the importance of face validity assessments across various fields, including healthcare, environmental economics, and transportation. Future research should continue to refine these assessments, supporting meaningful public engagement in policy and decision-making processes.

3.3 Hypotheses on PVE Evaluation

Prior to analyzing the outcomes of the data analysis in the upcoming chapter, several hypotheses are formulated based on existing literature. This study presents a total of five hypotheses about the evaluation of PVE consultations.

H1: Older individuals tend to evaluate PVE more positively compared to younger individuals.

Older individuals tend to evaluate PVE more positively compared to younger individuals. Research suggests that health conditions may deter older adults from active in-person participation; however, this barrier is mitigated in PVE as it is conducted online, resembling online surveys (Aini et al., 2017; Frogren et al., 2022). Studies indicate that younger individuals show the lowest response rates to both Internet and mail surveys. In contrast, seniors exhibit the highest response rates to Internet surveys, implying a greater willingness to engage in online participatory activities (Gigliotti & Dietsch, 2014). This heightened responsiveness among older individuals may be attributed to their flexibility and more availability of times. Additionally, it may be caused of their positive outlook on public participation, as they are more likely to believe in their ability to influence municipal administrations and decision-making processes (Avcıoğlu, 2023). This sense of political efficacy, where older individuals feel their actions can make a difference, contrasts sharply with the feelings of alienation and disbelief in the effectiveness of political participation often expressed by younger individuals (Avcıoğlu, 2023).

H2: Women generally have a positive attitude towards PVE compared to men

It is hypothesized that women will generally exhibit a more positive attitude towards PVE compared to men. Research by Wester and Morn (2013) indicates that a higher proportion of women than men express a desire for more participatory processes, even if they may not actively participate themselves. This aligns with the findings of Coffé and Bolzendahl (2009), who argue that women are more inclined to engage in private forms of activism, such as signing petitions and boycotting products for political reasons. This behavior suggests a preference for participatory actions that can be seamlessly integrated into daily life, particularly those that involve informal and private political engagement. Additionally, studies demonstrate that women and older individuals are generally more willing to participate in research surveys compared to younger men, further supporting the notion that women are more receptive to participatory methods like PVE, which often employ survey techniques (Glass et al., 2015).

H3: Highly educated respondents view PVE positively but with a more critical perspective.

Higher education levels are associated with increased participation in environmental impact assessments and public policies. Educated individuals possess better literacy and communication skills, enabling them to provide valuable input and critically assess political systems and policy implementations (Carreira et al., 2016; Onyango et al., 2019).

H4: Frequent users of the train or specific project tend to have a more positive evaluation of PVE.

It is hypothesized that frequent users of the train or a specific project will tend to have a more positive evaluation of PVE. Research by Harris et al. (2018) suggests that familiarity with the subject of study can positively influence respondents' perceptions, leading to more engaged and open responses. Therefore, it is predicted that frequent train travelers, who value service aspects such as speed, punctuality, safety, comfort, and customer care, are likely to have a positive attitude towards participating in studies about trains and mobility, which are the main topics of PVE consultations utilized in this study.

H5: Individuals living closer to the project are more likely to have a positive perception of PVE.

Proximity to a project or technology implementation site can lead to a more positive evaluation due to perceived benefits outweighing risks. This effect is observed in communities near heritage sites, where proximity enhances perceived value and increases engagement in conservation efforts (Huijts et al., 2019; Oladeji et al., 2022). Consequently, it is postulated that those who reside in proximity to the project being examined are more likely to hold favorable opinions of the PVE consultations.

Table 3.1: Current State of Knowledge of Face Validity Assessment

No	Source	Instrument being Assessed	Object Assessed by Face Validity
1	Glenk et al., 2024	Choice experiments for environmental and agricultural valuation	The degree to which WTP estimates are plausible
2	Dalawi et al., 2023	Questionnaire on COVID-19 knowledge and practices	Ease of comprehension and clarity of the questionnaire
3	Öztürk, 2023	Large-scale simulation projects	Model's credibility based on expert experience
4	Morton et al., 2023	Evaluation Tool of Health Information for Consumers (ETHIC)	Perceived usability and comprehensibility of the tools
5	Golan (2023)	PVE in Tel Aviv	Participants' perceptions of the instrument's suitability, perceived fulfilling its intended purpose and relevancy
6	Banna et al., 2023	A physical activity questionnaire	Comprehensiveness, relevance, and suitability of the questionnaire
7	Engel et al., 2023	Four preference-weighted quality-of-life instruments	Comprehensiveness of the questions and their appropriateness in the context of aged care
8	Hadi et al., 2023	A questionnaire assessing financial risk tolerance	Comprehensiveness and relevance of the tools
9	Juschten and Omann, 2023	Participatory Value Evaluation (PVE) for transport policy decisions	How well PVE reflects public opinion and decision-making processes in policy evaluations
10	Reffien et al., 2022	An evaluation checklist for the FDC program	Comprehensiveness, clarity, and relevance of the checklist items
11	Carlton et al., 2022	EQ Health and Wellbeing (EQ-HWB) measure	Appropriateness, relevancy, and comprehensiveness of the items in the questionnaire
12	Tuit, 2022	Face validity statements attached to PVE consultation	Participants' perceptions of the instrument's appropriateness and effectiveness in capturing their views on policy options
13	Babalola et al., 2022	A quality-of-care assessment tool for community health workers	Clarity, comprehensiveness, and acceptability of the tools
14	Mason et al., 2020	The AV User Perception Survey	Whether the survey items appeared credible and understandable to laypersons
15	Patel and Desai, 2020	A Questionnaire	Measure the clarity, unambiguity, reasonability, and relevancy of the questionnaire
16	Connell et al., 2018	Recovering Quality of Life (ReQoL) measure	How well items reflect user experiences and are understood by service users
17	Taherdoost, 2016	A questionnaire or survey used for data collection	The appearance of the questionnaire in terms of feasibility, readability, consistency of style and formatting, and the clarity of the language used
18	Hermans et al., 2016	interRAI Palliative Care instrument	The extent to which the instrument's items accurately reflect the needs and conditions of palliative care patients
19	Nevo, 1985	Concept of face validity	How well the test aligns with its stated objectives and whether it seems appropriate and relevant for its intended use

3.4 Summary

Literature Review

This chapter provided answers to **sub-question 1.1**, *"What is the current state of knowledge in literature, regarding face validity in public participation approaches, particularly within the realm of transportation planning?"*.

- **Face Validity**, one of the quickest methods for validating tools, assesses whether an instrument appears to measure what it is intended to. Despite criticisms of its subjectivity, face validity remains crucial in public participation methods like PVE, as it helps ensure that participants perceive the process as legitimate and relevant.
- The **literature review** on current knowledge of face validity in public participation tools reveals the application of face validity across various fields, including healthcare, environmental economics, and public participation. In each field, ensuring that instruments are clear and relevant is particularly important for the target audience or respondents.
- PVE has been evaluated for face validity, with a focus on clarity, relevance, and comprehensibility. However, there is a lack of studies examining the specific groups of citizens who consistently or negatively evaluate the face validity of PVE, as well as the reasons behind these evaluations.
- Five **hypotheses** are proposed to explore how different demographic groups perceive the face validity of PVE. These hypotheses consider the potential influence of factors such as age, gender, education, frequency of use, and proximity on participants' evaluations.

4 Evaluation of Face Validity in PVE Consultations

This chapter presents the findings of the quantitative data analysis that addresses the second sub-question: *"What are the characteristics of the distinct segments in PVE consultations, and how do these characteristics relate to their face validity evaluations?"* Additionally, it examines the consistency of these face validity evaluations across different PVE experiments in transportation planning, which is the focus of the third sub-question.

The chapter begins by examining the characteristics of the sample from all three consultations, and next presents a descriptive analysis of the PVE consultations evaluation. Subsequently, the outcomes of cross-tabulation analysis are presented to determine any relationships between the respondents' attributes and their assessments of face validity. Following that, the process of identifying segments within the sample is examined using Latent Class Cluster Analysis. The study investigates the influences of respondents' attributes and their ratings of face validity using Multinomial Logistic Regression. At the conclusion of each analysis, a comparison is made between the three consultations.

4.1 Sample Characteristics

This section presents a summary of the main characteristics of the participants in the three consultations: Lelylijn, Oude Lijn, and Mobility Vision. The subsections provide detailed information about the consultation respondents, which is obtained from a dataset that has been factor-weighted. The procedure of assigning weights to factors and its consequences for data analysis are explained in the appendix B.

4.1.1 Lelylijn Consultation

The main characteristics of respondents in the Lelylijn consultation are outlined in Table 4.1. This characteristic was derived from factor-weighting the actual dataset, a process used to assure the generalizability of the data analysis conclusion. This process involved assigning more weight to responses from participant groups underrepresented in the sample, while reducing the weight of over-represented groups using a weighting factor, as detailed in Appendix B. The re-weighting process eventually results in total 4.572 respondents in the sample of Lelylijn project. Furthermore, the statistical analysis of the weighted dataset yielded better outcomes in terms of interpretability and statistics, than the non-weighted dataset's results in the Multinomial Logistics Regression Analysis results and Latent Class Cluster Analysis. Thus, the weighted dataset is subsequently utilized in further quantitative analysis.

Table 4.1: Lelylijn Project Consultation's Sample Characteristics used for Further Quantitative Analysis. The sum in each attribute might not be equal due to factor-weighting process.

Respondents' Attribute	Total Respondents*	Respondents %	Population Netherlands %
Gender			
Man	2.332	51%	51%
Woman	2.240	49%	49%
Age			
18-34 years	1.215	27%	27%
35-64 years	2.220	49%	48%
65 years or older	1.137	25%	25%
Education			
Low	1.245	27%	29%
Medium	1.684	37%	37%
High	1.643	36%	35%

In addition to these main demographic characteristics, there are other respondents' characteristics that are part of this research. An overview of all main and case-specific characteristics and their categorization is presented in Appendix B.1.

4.1.2 Oude Lijn Consultation

The main characteristics of respondents in the Oude Lijn consultation are described in Table 4.2. Similar to Lelylijn dataset, these characteristics'proportions were derived from factor-weighting the actual dataset to ensure the generalizability of the data analysis conclusions. This approach is similar to the method used in the Lelylijn consultation, as described in Appendix B.1.2. The initial dataset comprised 3.844 respondents, and after the factor-weighting process, the sample size for the Oude Lijn project increased to 4.070 respondents. The weighted dataset was subsequently used for further analysis of the Oude Lijn consultation. A comprehensive overview of all main and case-specific characteristics in Oude Lijn consultation and their categorization is provided in Appendix B.2.

Table 4.2: Oude Lijn Project Consultation's Sample Characteristics used for Further Quantitative Analysis. The sum in each attribute might not be equal due to factor-weighting process.

Respondents' Attribute	Total Respondents*	Respondents %	Population Netherlands %
Gender			
Man	2.049	50%	51%
Woman	2.021	50%	49%
Age			
18-34 years	1.118	27%	27%
35-64 years	1.964	48%	48%
65 years or older	987	25%	25%
Education			
Low	1.115	27%	29%
Medium	1.508	37%	37%
High	1.447	36%	35%

4.1.3 Mobility Vision Consultation

The main characteristics of respondents in the Mobility Vision consultation are described in Table 4.3. These characteristics, derived from factor-weighting the actual dataset, ensure the generalizability of the data analysis conclusions. This approach is similar to the method used in the Lelylijn consultation, as described in Appendix B.1.2. The initial dataset comprised 4.146 respondents, and after the factor-

weighting process, the sample size for the Mobility Vision project increased to 4.258 respondents. The weighted dataset was subsequently used for further analysis of the Mobility Vision consultation. A comprehensive overview of all main and case-specific characteristics in Mobility Vision consultation and their categorization is provided in Appendix B.3.

Table 4.3: Mobility Vision Project Consultation's Sample Characteristics used for Further Quantitative Analysis. The sum in each attribute might not be equal due to factor-weighting process.

Respondents' Attribute	Total Respondents*	Respondents %	Population Netherlands %
Gender			
Man	2.161	51%	51%
Woman	2.097	49%	49%
Age			
18-34 years	1.134	27%	27%
35-64 years	2.041	48%	48%
65 years or older	1.083	25%	25%
Education			
Low	1.190	28%	29%
Medium	1.565	37%	37%
High	1.502	35%	35%

4.2 Descriptive Results

Descriptive statistics were employed to evaluate the ratings given by the respondents regarding the overall and the face validity of PVE consultations. Each of the three consultations consisted of seven face validity statements that required responses in a form of Likert scale. Although the face validity claims in the three consultations were phrased in somewhat different ways, the primary content of each statement remained largely the same. Table 4.4 displays a comparison of the presence of statements in each consultation. The subsequent sections provide the outcomes of the descriptive analysis from each consultation.

Table 4.4: Summary of Presence of Face Validity Statements in Three Consultations, based on Figures 2.1, 2.3a, and 2.4a.

No	Face Validity Dimension	Statement	Consultation		
			Lelylijn	Oudelijn	Mobility Vision
1	Transparency	The consultation steered my choices in a certain direction	✓	✗	✗
2	Transparency	I trust this is an honest/fair investigation	✓	✓	✓
3	Feasibility	I found the consultation difficult to understand	✗	✓	✓
4	Relevance	I thought it was an important topic to give my opinion	✓	✓	✓
5	Relevance	By participating in this consultation, I have learned about the choices that the government has to make on this subject	✓	✓	✓
6	Acceptance	This method should be used more often to involve residents in public policy	✓	✓	✓
7	Acceptance	If many people participate in this consultation, then the final decisions on this subject are better accepted for me	✓	✓	✓
8	Acceptance	If the government often allows residents to think along about these types of choices more often, then I get more confidence in the decisions of the government	✓	✓	✓

4.2.1 Lelylijn Consultation

Figure 4.1 and Appendix C.1 presents the results of face validity evaluations conducted in the Lelylijn consultation. There is overall positive evaluation of Lelylijn consultation. 37,4% of the sample responded neutrally to the negatively framed statement "The consultation steered my decisions in a specific direction", while 46,8% of the sample strongly disagreed with the statement. The strong disagreement expressed by almost half of the respondents indicates that the majority of participants do not believe that the Lelylijn consultation steered their decisions. However, it is important to consider that a significant proportion (15,9%) of respondents still believe that the Lelylijn consultation is steering their choices. Consequently, this fact should be taken into account in future PVE designs. The factors contributing to these adverse perceptions are further examined in chapter 5.

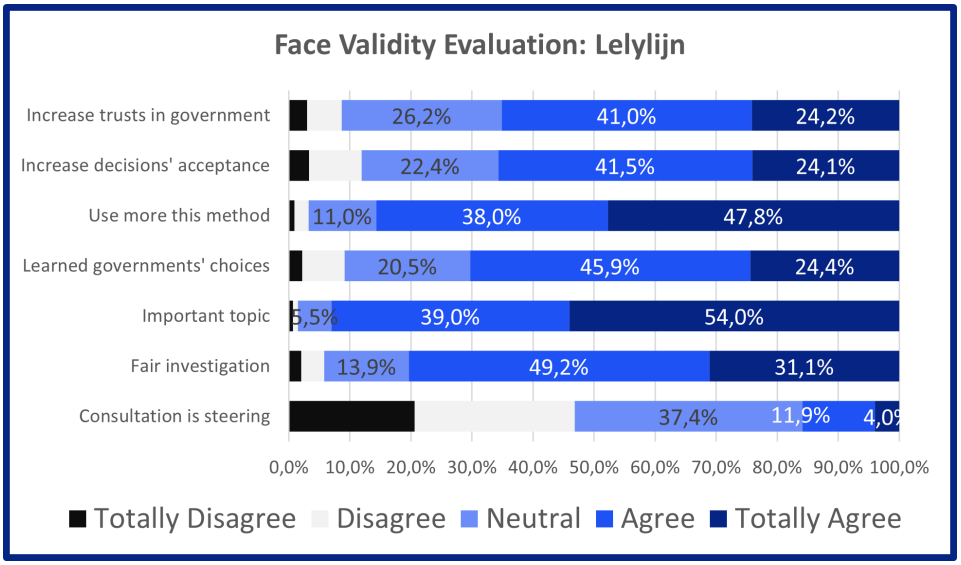


Figure 4.1: Face Validity Evaluation in Lelylijn Dataset

On the other side, positively formulated statements towards Lelylijn consultation mostly have (totally) agreed assessment by the sample. 80,3% of the sample (totally) agree that this consultation is a fair investigation. 93% of the sample think that Lelylijn project was an important topic to give their opinion on. 85,7% respondents think that this consultation should be used more often to involve residents in government policy. They also think that by participating in this consultation, they learn about the choice the government has to make about the Lelylijn (70,3%). The respondents think that the residents involvement on large scale through this consultation will increase their acceptance on the final decisions of about the Lelylijn (65,6%). Finally, 65,1% of respondents (totally) agree that if the government allows them to think about these types of choices more often, it will increase their trust in government decisions.

4.2.2 Oude Lijn Consultation

A descriptive analysis was performed on the face validity evaluation in the Oude Lijn consultation dataset. Figure 4.2 and Appendix C.3 presents the overall positive evaluation of face validity by the respondents.

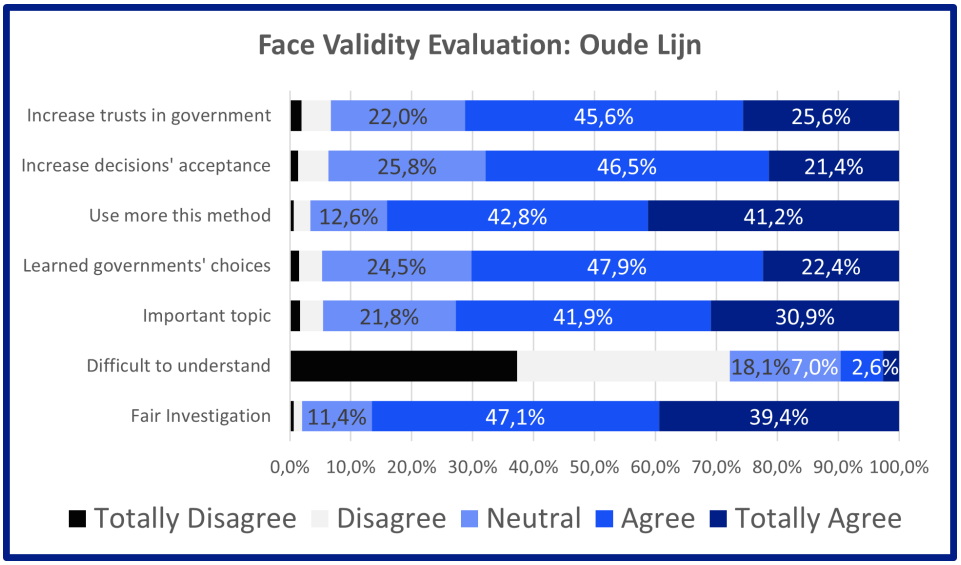


Figure 4.2: Face Validity Evaluation in Oude Lijn Dataset

Regarding the transparency dimension, a significant 86,6% of respondents (totally) agreed that this consultation is an fair investigation. This statement also has the highest level of agreement across other positively formulated statements in the Oude Lijn dataset.

In terms of feasibility dimension, the statement was intentionally formulated negatively to assess the respondents' consistency and focus. For statement "I found the consultation difficult to understand", the analysis revealed that 72,2% of respondents (totally) disagreed with this statement, indicating that the Oude Lijn consultation was generally perceived as feasible and understandable. This aligns with the overall positive evaluations indicated by the high levels of agreement with subsequent positively formulated statements, suggesting that respondents maintained focus throughout the consultation. Nevertheless, it is worth mentioning that 18,1% of the participants expressed a neutral opinion, while 9,6% fully agreed that the consultation was challenging to understand. This suggests the necessity of further improvements in PVE consultations.

Regarding relevance dimension, two statements were posed in the Oude Lijn consultation. First, 72,8% of respondents (totally) agreed that Oude Lijn was an important topic for them to provide their opinions on. Additionally, 70,2% (totally) agreed that by participating in the consultation, they learned about the choices that the government must make on this subject. Only 5,4% of respondents (totally) disagreed that the consultation was important, and 5,2% (totally) disagreed that they learned about the government's decision-making choices. These findings suggest that the Oude Lijn consultation was relevant to the respondents, as they viewed the topic as important and felt that their participation helped them understand the government's challenges in decision-making on the topic.

Finally, in the acceptance dimension, three statements were evaluated. First, 84,1% of respondents (totally) agreed that this method of consultation should be used more frequently to involve residents in public policy, indicating an overall acceptance regarding of PVE method. Secondly, 67,9% (totally) agreed that if many people participate in the consultation, it will increase their acceptance on the final decisions on the subject. Lastly, 71,3% (totally) agreed that if the government frequently allows residents to think along with these types of choices, they will experience that participating in PVE increases their trust in the government. However, a notable portion of respondents (totally) disagreed with the latter two statements. These results indicate that while respondents generally accept the consultation, a number of participants still struggle to increase their acceptance in the decisions and trust in the government, even with opportunities for citizen participation.

4.2.3 Mobility Vision Consultation

A descriptive analysis was performed on the face validity evaluation in the Mobility Vision consultation dataset. Figure 4.3 and Appendix C.5 presents face validity evaluations that cover seven statements

across different face validity dimensions: Transparency, Feasibility, Relevance, and Acceptance.

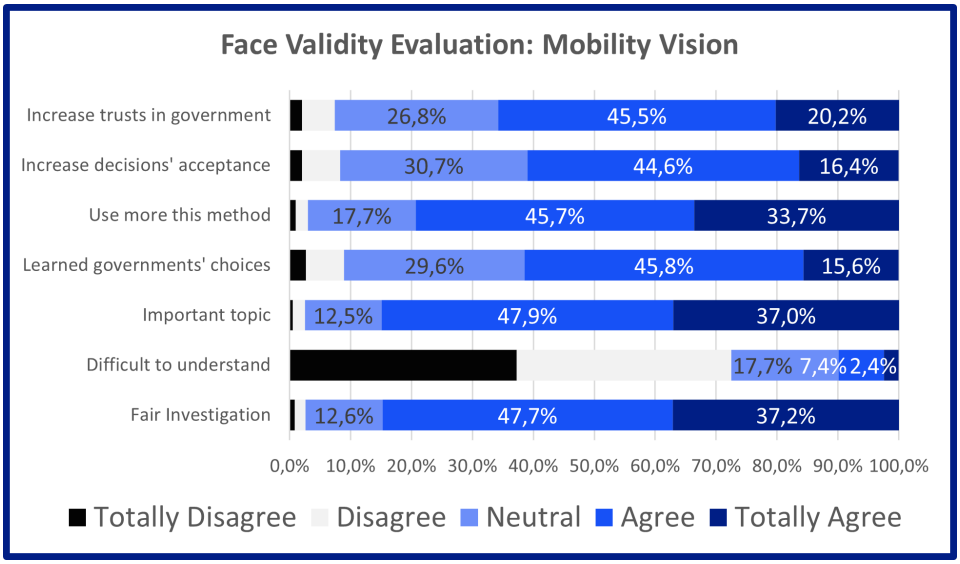


Figure 4.3: Face Validity Evaluation in Mobility Vision Dataset

For the first statement under the Transparency dimension, ”I trust that this research project is conducted in a fair way,” 84,9% of respondents (totally) agreed, indicating high trust and consistently positive responses. Next, under the Feasibility dimension, a majority of respondents (72,4%) (totally) disagreed with the second statement, ”I found the consultation difficult to understand”, suggesting most participants did not find the consultation difficult. Nevertheless, similar to Oude Lijn consultation, there is notable proportion of respondents who were neutral (17,7%) and (totally) agreed that the consultation was difficult to understand (9,8%). This observation will be further discussed in the discussion section.

In terms of the Relevance dimension, the third statement, ”I believed it was an important topic to give my opinion on,” garnered 84,9% (total) agreement, suggesting that Mobility Vision is a significant and important subject for the majority of respondents. The fourth statement, which states that participants gained knowledge about the decisions the government needs to make on this matter, received a total agreement rate of 61,4% and 29,6% of respondents remained neutral. While a substantial majority of the respondents believed they gained knowledge through this consultation, there is a notable percentage (9,0%) of respondents who felt that they did not learn about the government’s decisions through this consultation. Hence, it is imperative to discuss and address this matter in subsequent sections, as one of the primary purposes of public involvement methods like PVE is to serve as effective two-way communication tools between respondents and governments (Quick, 2014).

Within the Acceptance dimension, the fifth statement, which inquires for increased utilization of this method to engage residents in government policy, had significant positive response with 79,4% (total) agreement or complete agreement. This suggests a strong approval and consistent answers from the participants. The sixth statement, ”If many people participate in this consultation, the final decisions on this topic are more acceptable to me”, garnered a 61,0% (total) agreement. However, 30,7% of the respondents expressed neutrality towards the statement, suggesting that a significant portion of respondents will unlikely to increase their acceptance on the final decisions about Mobility Vision, despite greater participation in this consultation.

Finally, the statement ”If the government involves residents more often in thinking about these kinds of choices, I will have more confidence in the government’s decisions”, obtained 65,7% (total) agreement. Similar to the previous statement about government’s decisions’ acceptance, there are notable amount of neutral (26,8%) and disagreement (7,4%) with the statement. This indicates the need for further investigation into the causes behind these responses and to address the low likeliness to increase trust in the government within the sample.

4.2.4 Comparison of Face Validity Evaluation Across Consultations

This section will examine the comparison between the descriptive analysis of face validity evaluation and the overall consultation assessment across the three PVE consultations. The comparison will be examined based on the dimension of face validity, namely transparency, feasibility, relevance, and acceptance.

Transparency & Feasibility Dimension

The results of the descriptive statistics analysis for each consultation indicated a generally positive appraisal of face validity in all three PVE consultations. According to the analysis of negatively formulated statements shown in Figure 4.4 and 4.5, about 46,8% of respondents did not think that the consultation influenced their choices. Meanwhile, over 70% of respondents find this consultation to be understandable, in both the Oude Lijn and Mobility Vision consultations. However, a notable proportion of respondents remain impartial (37,4%) and consider the Lelylijn consultation as steering their decisions (15,9%). Additionally, approximately 9,7% of respondents in Oude Lijn and Mobility Vision think that the consultation is difficult to understand. This is a concern that needs careful consideration and should be taken carefully when it comes to improving future PVE designs.

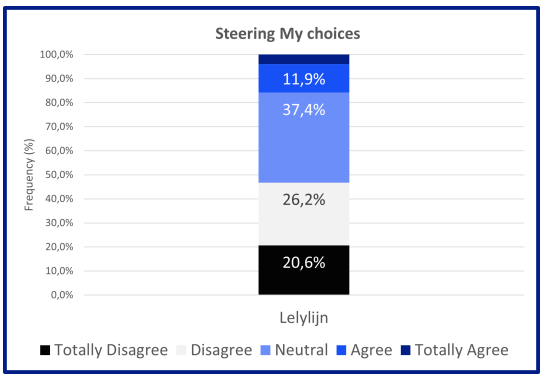


Figure 4.4: This Consultation is steering

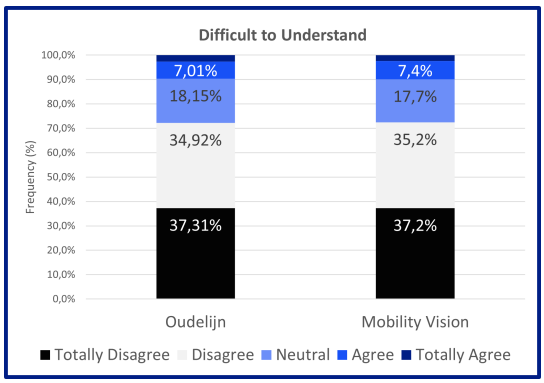


Figure 4.5: This Consultation is Difficult to Understand

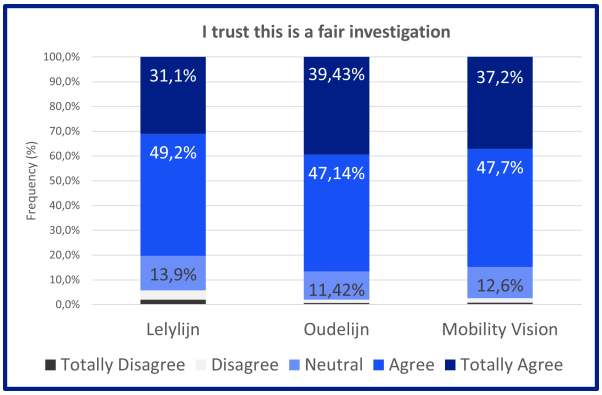


Figure 4.6: Trust Investigation is Fair

Figure 4.7: Face Validity Evaluation in Transparency and Feasibility Dimensions

Another statement regarding transparency is ”I trust this is a fair investigation,” which was included in all three consultations. Across all three consultations, as depicted in Figure 4.6 over 80% of the respondents expressed confidence in the integrity of the consultation. Remarkably, less than 3% of the respondents in Oude Lijn and Mobility Vision believe that this consultation lacks honesty. Although slightly higher, solely 5,8% of respondents in the Lelylijn consultation believe that this consultation is not truthful.

Relevance Dimension

Within the relevance dimension, as depicted in Figure 4.10, there are two statements that are present in all three PVE consultations. According to the statement, 93% of the respondents of Lelylijn believe

that the consultation is important, considering it to be a major topic to express their view on. One probable explanation is that the Lelylijn has not yet been put into operation, prompting the respondents to consider it essential to express their views. The finding is noteworthy due to the fact that a majority of the participants (54,3%) in the Lelylijn consultation live far from the project. The Oude Lijn consultation, although recognized to be important by a large proportion of individuals, had the least consensus regarding its importance (72.8%) compared to the other two consultations. One possible reason for this is because although the consultation asked for suggestions for improvements and changes, the Oude Lijn already exists and is now able to provide ”sufficient” services to the respondents.

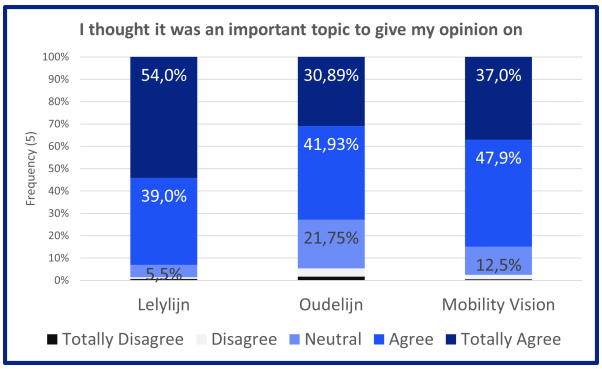


Figure 4.8: Important Topic

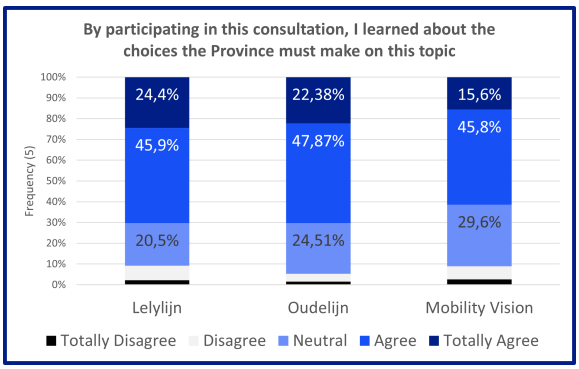


Figure 4.9: Learned about Government's Choices

Figure 4.10: Face Validity Evaluation in Relevance Dimension

Another statement in the Relevance dimension is ”By participating in this consultation, I have learned about the choices that the government has to make on this subject”. Both the Lelylijn and Oude Lijn consultation indicate that 70,3% of its respondents believe that they learned about the government's decisions through the consultation. In the Mobility Vision consultation, a lower portion of respondents of 61,4%, (totally) agreed with the statement. However, approximately 9,0% of the participants in the Lelylijn and Mobility Vision surveys expressed disagreement with the statement, while at least 20,0% of the participants remained neutral in all three consultations. This indicates that a significant percentage of participants hold the belief that they did not receive sufficient education regarding the decisions that the government must make regarding the topic. As a result, there is a possibility for future PVE designs to improve the the ability of PVE to inform the citizens about the government's actions through consultations.

Acceptance Dimension

Within the acceptance dimension, depicted in Figure 4.14, there are three face validity statements that are included in all three PVE consultations. A minimum of 79% of the participants in three consultations (totally) agreed with the initial statement ”This method should be used more often to involve residents in government policies”. A mere 3,4% of the respondents expressed their disagreement with using this strategy more frequently. This indicates that respondents generally believe PVE consultation should be utilized more frequently to engage residents in government policies.

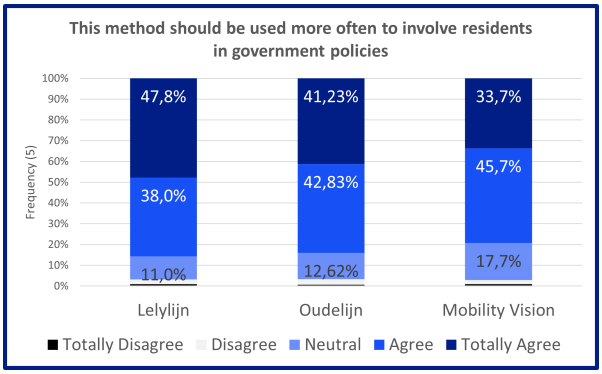


Figure 4.11: This Method Should be Used More

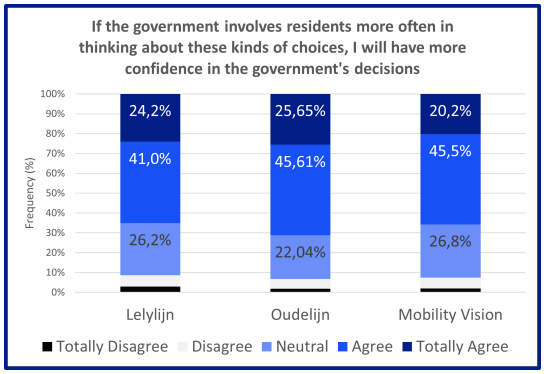


Figure 4.12: Final Decisions are More Acceptable

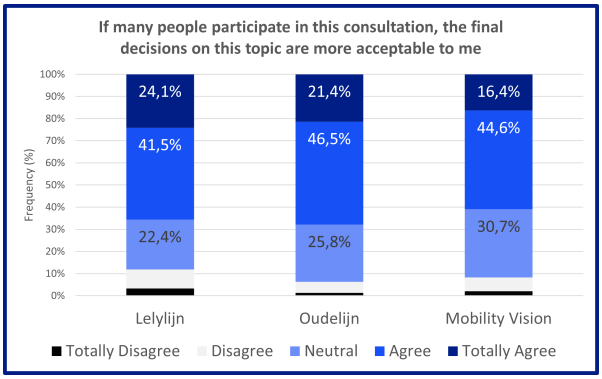


Figure 4.13: Confidence in Government's Decisions

Figure 4.14: Face Validity Evaluation in Acceptance Dimension

The second statement, "If many people participate in this consultation, the final decisions on this topic are more acceptable to me," also elicits a significant level of agreement, with at least 61% of the respondents throughout the three consultations expressing agreement. Nevertheless, a minimum of 6,6% of the participants express (total) disagreement, while at least 22,4% remain indifferent, throughout all three consultations. This suggests that while the majority of respondents think that participating in PVE will increase their acceptance on the final decisions on the subject, there is still a subset of individuals who are unlikely to increase their acceptance on the final decisions, regardless of the high number of participants in the consultation. However, this number could also be because their acceptance on the final decisions were already high regardless of their participation in PVE.

The final statement, "If the government involves residents more often in thinking about these kinds of choices, I will have more confidence in the government's decisions" exhibits similar patterns to the preceding statement regarding public receptiveness towards final decisions. Over 65,1% of the participants in all three consultations experience that participating in PVE increases their trust in the government. Conversely, a minimum of 6,7% of the participants express skepticism regarding their future trust in the government

Consultation Ratings

The overall rating of the consultation was inquired to respondents in each consultation. The respondents' assessments average scores of 7,42; 7,72; and 7,55 for the Lelylijn, Oude Lijn, and Mobility Vision consultations, respectively. The ratings distribution, as depicted in Figure 4.15, indicates that although the distributions across the three consultations appear identical, the Oude Lijn consultation tends to surpass the other two consultations in terms of higher ratings. This suggests that a greater number of individuals rate Oude Lijn consultation more favorably compared to the others. Lelylijn has then lowest average ratings among the 3 PVE consultations. Furthermore, there is a minority of individuals who score the Lelylijn consultation as 1 and 3, although the number is not substantial. The more negative evaluation of Lelylijn will be subject to additional examination in the subsequent Chapter 5.

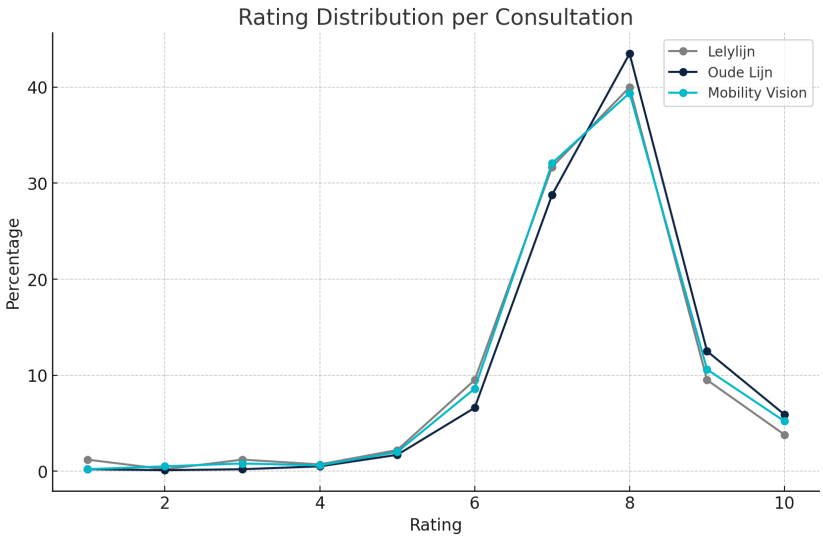


Figure 4.15: Consultation Ratings Distribution Across 3 PVE Consultations

To summarize, the evaluation of face validity statements and the assessment of overall ratings throughout three PVE consultations generally indicate positive evaluation. There was a greater number of positive ratings, particularly in relation to the significance of the important of topic and the fairness of the investigation. Nevertheless, according to the assessment, there is room for enhancing PVE designs in terms of the respondent's trust in the government's final decisions and making the consultation less steering. The comparable descriptive outcomes observed across the three datasets suggest a level of consistency that strengthens the reliability of the findings, further justifying the utilization of these three datasets in this study.

4.3 Crosstabulation Analysis Results

In this section, crosstabulation analysis results are discussed to provide a descriptive study of the association between respondents' characteristics with their PVE evaluation. The associations observed in this study pertain to trends within the sample and should not be used to make predictive or inferential judgments about the broader population. Nevertheless, the presence of statistically significant results indicates that the observed interactions are unlikely to have occurred by chance. This suggests that similar relationships may also be present in the broader population from which the sample was selected. The presence of non-significant parameters suggests that a meaningful association between the variables cannot be inferred.

The following sections present the results of the crosstabulation analysis for overall consultation ratings in all PVE consultations. This is followed by a detailed analysis of the crosstabulation results for the evaluation of face validity statements in each PVE consultation.

4.3.1 Crosstabulation Analysis on Overall Ratings

Crosstabulation analysis of the respondents' characteristics with how they rate the overall PVE consultations were performed to get the preliminary insights on whether there are significant relationships between certain characteristics of respondents with how they rate the consultation experience in general. Table 4.16 presents the results of the crosstabulation analysis across 3 PVE consultations. Each attribute or respondents' characteristic has average ratings per level and per consultation. The Gamma coefficient, which ranges from -1 to 1, measures the extent of the connection, with an absolute value of 1 indicating a strong correlation. A positive value of Gamma signifies a direct relationship between the respondents' characteristic with their consultation ratings. As the level of the characteristics increases, their ratings also tends to increase. A significance (Sig.) value below 0,05 implies that the observed correlation is statistically significant, indicated by green cells. The empty cells indicate that the particular attribute was not present in the consultation.

Attribute	Level	Consultation								
		Lelylijn			Oudelijn			Mobility Vision		
		AVG Ratings	Gamma	Sig	AVG Ratings	Gamma	Sig	AVG Ratings	Gamma	Sig
Age	18-34 years	7,53	-0,143	<,001	7,78	-0,037	0,073	7,53	0,036	0,074
	35-64 years	7,28			7,69			7,52		
	65 years or older	7,42			7,65			7,65		
Gender	Man	7,39	-0,009	0,706	7,77	-0,081	0,001	7,60	-0,099	<,001
	Woman	7,45			7,67			7,50		
Education	Low	7,41	0,081	<,001	7,73	0,001	0,969	7,43	0,016	0,432
	Medium	7,42			7,70			7,61		
	High	7,42			7,75			7,57		
Financial	Not enough	7,47	-0,024	0,325	-	-	-	-	-	-
	Enough	7,43								
	More than enough	7,37								
Occupation	Not working	7,39	-0,049	0,01	7,72	0,025	0,226	7,52	0,012	0,624
	Part-time	7,45			7,62					
	Full-time	7,44			7,79			7,57		
Residence	Tenant	7,55	-0,154	<,001	7,72	-0,012	0,587	-	-	-
	Lodger	7,43			7,75					
	Homeowner	7,37			7,71					
Proximity	Not proximate	7,32	0,078	<,001	7,67	0,058	0,02	-	-	-
	Proximate	7,53			7,77					
Freq Use Train	Rarely	7,25	0,173	0	7,58	0,173	<,001	-	-	-
	Occassionally	7,73								
	(Almost) daily	7,53			7,85					
Freq Use Line	Rarely	7,41	0,063	0,002	7,62	0,174	<,001	-	-	-
	Occassionally	7,43								
	(Almost) daily	7,42			7,89					
In/out from at station	In/ out at any station	-	-	-	7,91	-0,192	<,001	-	-	-
	Never in/out at any station				7,61					
Work or Live near station	Live/work at any station	-	-	-	7,89	-0,179	<,001	-	-	-
	Never Live/work at any statior				7,62					
Type of area of living	Urban	-	-	-	7,76	-0,082	0,002	7,58	-0,035	0,155
	Rural				7,62			7,50		
Freq Problems	Low	-	-	-	-	-	-	7,61	-0,114	<,001
	Moderate							7,47		
	High							7,35		

Figure 4.16: Crosstabulation Analyisis of Respondents' Characteristics and Their Consultation Ratings

The Lelylijn dataset shows a slight negative correlation between age, with a Gamma value of -0,143 and a p-value of less than 0,001. This indicates that older participants are more inclined to provide lower ratings for the consultation. With respect to the gender attribute, there are two consistent despite weak significant associations between an individual's gender and their consultation evaluations (Gamma = -0,081, p = 0,001 in Oude Lijn; Gamma = -0,099, p < 0,001 in Mobility Vision). This implies that women have a lower probability of giving a higher rating to the consultation.

The respondent's education level relationship with how they rate the consultation is only significant in the Lelylijn dataset (Gamma = 0,081, p < 0,001). It is evident that those with a greater degree of education are more inclined to give a higher rating to the consultation. There is a weak correlation between occupation (Gamma = -0,049, p = 0,01) and residence (Gamma = -0,154, p < 0,001). These data indicate that those who are unemployed and reside in rented housing are more inclined to give a higher rating to the consultation.

The proximity of an individual to the project has a weak and consistent correlation with their rating of the consultations. In the case of Lelylijn, the correlation coefficient (Gamma) is 0,078, with a p-value of less than 0,001. In the case of Oude Lijn, the correlation coefficient is 0,058, with a p-value of 0,02. This suggests that individuals residing in closer proximity to the project are more inclined to give a positive rating to the consultation. Similarly, there appears to be a positive correlation between the frequency

of train usage and the likelihood of rating the consultation favorably.

The research also indicates that those who have used the stations in the study are more likely to give the consultation a higher grade (Gamma = -0,192, p < 0,001). Likewise, individuals who work or reside in close proximity to the examined stations are more inclined to give a higher rating to the consultation (Gamma = -0,179, p < 0,001). Ultimately, individuals residing in urban areas (Gamma = -0,082, p = 0,002) and those who rarely encounter issues with accessibility (Gamma = -0,114, p < 0,001) are also inclined to assign a higher rating to the consultation.

4.3.2 Lelylijn Consultation

Crosstabulation analysis was conducted for all statements in the Lelylijn consultation. The findings related to the transparency and relevance dimension are presented in Table 4.5, whilst the results of the acceptance dimension are provided in Table 4.6. The green cells indicate significant association. For the crosstabulation analysis on the assessments of face validity statements, the options were classified into three tiers: disagree, neutral, and agree. The actual responses labeled as ”totally disagree” and ”disagree” are grouped together in this analysis under the category of ”disagree,” while the responses labeled as ”totally agree” and ”agree” are grouped together under the category of ”agree”. The detailed interpretation will be further examined for each face validity statement below.

The consultation steered my choices in a certain direction

There is a weak positive association between age and the belief that the consultation steered choices (Gamma = 0,050, p = 0,021), suggesting that older respondents are slightly more likely to feel that the consultation influenced their choices. A moderate negative association exists between gender and this belief (Gamma = -0,187, p < 0,001), indicating that women are less likely to feel that the consultation steered their choices compared to men.

Additionally, a moderate negative association is observed between education level and this belief (Gamma = -0,145, p < 0,001), implying that respondents with higher education levels are less likely to feel influenced by the consultation. Financial status also shows a moderate negative association (Gamma = -0,136, p < 0,001), with those having more financial resources every month being less likely to feel that their choices were steered by the consultation.

Occupation has a weak negative association with this belief (Gamma = -0,092, p < 0,001), indicating that respondents who work full-time are slightly less likely to feel influenced by the consultation. Proximity to the Lelylijn shows a weak positive association (Gamma = 0,089, p < 0,001), suggesting that respondents living closer to the Lelylijn are slightly more likely to feel that their choices were influenced. Lastly, there is a weak negative association between the frequency of using trains and the belief that the consultation steered choices (Gamma = -0,117, p < 0,001), indicating that those who use trains less frequently are slightly less likely to feel influenced by the consultation.

Table 4.5: Lelylijn Project Consultation's Cross-tabulation Analysis Results: Transparency and Relevance Dimension

Face Validity Dimension		Transparency						Relevance					
Statement		The consultation steered my choices in a certain direction			I trust this is a fair investigation			I thought it was an important topic to give my opinion on			By participating in this consultation, I learn about the choices the government has to make about the Lely Line		
Response		Disagree	Neutral	Agree	Disagree	Neutral	Agree	Disagree	Neutral	Agree	Disagree	Neutral	Agree
Gender	Man	21%	19%	10%	3%	7%	40%	1%	4%	45%	5%	12%	33%
	Woman	26%	19%	6%	3%	7%	40%	0%	2%	48%	4%	9%	37%
	Gamma	-0,187			0,021			0,370			0,144		
	Sig.	<0,001			0,558			<0,001			<0,001		
Age	18-34 years	23%	16%	7%	3%	4%	40%	1%	2%	44%	4%	7%	35%
	35-64 years	16%	12%	5%	2%	7%	25%	0%	2%	31%	3%	9%	22%
	65 years plus	8%	9%	3%	1%	4%	15%	0%	2%	18%	2%	5%	13%
	Gamma	0,050			-0,071			-0,278			-0,138		
	Sig.	0,021			<0,001			<0,001			<0,001		
Education	Low	9%	12%	4%	2%	2%	21%	0%	1%	23%	2%	5%	17%
	Medium	18%	14%	6%	2%	5%	30%	1%	2%	35%	3%	7%	27%
	High	20%	12%	6%	2%	6%	30%	1%	2%	35%	4%	8%	26%
	Gamma	-0,145			-0,034			-0,070			-0,049		
	Sig.	<0,001			<0,001			0,151			0,057		
Occupation	Not working	17%	17%	7%	2%	5%	34%	1%	2%	38%	4%	7%	30%
	Part-time	11%	8%	3%	1%	3%	17%	0%	1%	20%	2%	5%	15%
	Full-time	19%	13%	6%	3%	5%	29%	1%	2%	35%	4%	8%	25%
	Gamma	-0,092			-0,104			0,055			-0,097		
	Sig.	<0,001			<0,001			0,277			<0,001		
Residency	Tenant	12%	10%	5%	2%	3%	22%	0%	1%	26%	2%	4%	20%
	Lodger	7%	5%	2%	1%	1%	12%	0%	1%	13%	1%	3%	9%
	Homeowner	28%	22%	9%	4%	10%	47%	1%	4%	55%	6%	14%	41%
	Gamma	-0,023			-0,139			-0,293			-0,144		
	Sig.	0,340			<0,001			<0,001			<0,001		
Proximity	Not proximate	26%	18%	8%	3%	8%	41%	1%	3%	47%	5%	10%	36%
	Proximate	21%	19%	8%	3%	6%	40%	0%	2%	46%	4%	10%	34%
	Gamma	0,089			0,075			0,241			0,019		
	Sig.	<0,001			0,038			<0,001			0,532		
Frequency to use Train	Rarely	23%	22%	9%	4%	10%	40%	1%	3%	50%	5%	13%	36%
	Occassionally	12%	8%	3%	1%	2%	19%	0%	1%	21%	2%	3%	17%
	(Almost) Daily	537	322	176	39	93	903	10	41	984	88	176	771
	Gamma	-0,117			0,306			0,158			0,147		
	Sig.	<0,001			0,000			0,002			<0,001		
Frequency to use Lelylijn	Rarely	25%	19%	8%	3%	8%	41%	1%	4%	47%	5%	12%	36%
	Occassionally	14%	11%	4%	2%	4%	23%	0%	1%	27%	3%	6%	19%
	(Almost) Daily	370	300	178	45	79	725	3	17	829	56	142	651
	Gamma	0,043			0,142			0,394			0,078		
	Sig.	0,060			<0,001			<0,001			0,003		
Financial	Not enough	3%	2%	2%	0%	1%	6%	0%	0%	6%	0%	1%	5%
	Enough	31%	28%	11%	4%	11%	55%	1%	4%	65%	7%	15%	48%
	More than Enough	13%	8%	3%	2%	3%	20%	0%	1%	23%	2%	4%	18%
	Gamma	-0,136			-0,017			0,070			0,006		
	Sig.	<0,001			0,642			0,241			0,850		

I trust this is a fair investigation

There is a weak negative association between age and the belief that the investigation is fair (Gamma = -0,071, $p < 0,001$), suggesting that older respondents are slightly less likely to trust the investigation's fairness. Education level also shows a weak negative association with this belief (Gamma = -0,034, $p < 0,001$), indicating that respondents with higher education levels are less likely to trust the fairness of the investigation.

Occupation has a weak negative association with this belief (Gamma = -0,104, $p < 0,001$), showing that respondents who work full-time are slightly less likely to trust the investigation's fairness. Residency status reveals a moderate negative association (Gamma = -0,139, $p < 0,001$), with homeowners being less likely to trust the investigation compared to tenants and lodgers.

Proximity to the Lelylijn shows a weak positive association (Gamma = 0,075, $p = 0,038$), indicating that

respondents living closer to the Lelylijn are slightly more likely to trust the fairness of the investigation. The frequency of using trains has a moderate positive association with this belief (Gamma = 0,306, p < 0,001), suggesting that those who use trains more frequently are more likely to trust the investigation's fairness. Lastly, there is a weak positive association between the frequency of using the Lelylijn and the belief that the investigation is fair (Gamma = 0,142, p < 0,001), indicating that respondents who use the Lelylijn more often are more likely to trust the investigation.

I thought it was an important topic to give my opinion on

There is a moderate negative association between age and the belief that the topic is important (Gamma = -0,278, p < 0,001), suggesting that older respondents are less likely to consider it important to give their opinion on the topic compared to younger respondents. Gender shows a moderate positive association (Gamma = 0,370, p < 0,001), indicating that women are more likely to find the topic important to give their opinion on compared to men.

Residency status reveals a moderate negative association (Gamma = -0,293, p < 0,001), with homeowners being less likely to find the topic important compared to tenants and lodgers. Proximity to the Lelylijn has a moderate positive association (Gamma = 0,241, p < 0,001), indicating that respondents living closer to the Lelylijn are more likely to find the topic important to give their opinion on.

The frequency of using trains shows a weak positive association with this belief (Gamma = 0,158, p = 0,002), suggesting that those who use trains more frequently are more likely to find the topic important to give their opinion on. Similarly, the frequency of using the Lelylijn reveals a moderate positive association (Gamma = 0,394, p < 0,001), indicating that respondents who use the Lelylijn more often are more likely to find the topic important to give their opinion on.

By participating in this consultation, I learn about the choices the government has to make about the Lely Line

There is a weak negative association between age and the belief that participating in the consultation helped respondents learn about the government's choices regarding the Lely Line (Gamma = -0,138, p < 0,001), indicating that older respondents are slightly less likely to feel they learned about the government's choices compared to younger respondents. Gender shows a weak positive association (Gamma = 0,144, p < 0,001), suggesting that women are more likely to feel that they learned about the government's choices compared to men.

Occupation status reveals a weak negative association (Gamma = -0,097, p < 0,001), with full-time workers being less likely to feel that they learned about the government's choices compared to those not working or part-time workers. Residency status also shows a weak negative association (Gamma = -0,144, p < 0,001), indicating that homeowners are less likely to feel that they learned about the government's choices compared to tenants and lodgers.

The frequency of using trains has a weak positive association with this belief (Gamma = 0,147, p < 0,001), suggesting that those who use trains more frequently are more likely to feel that they learned about the government's choices. Additionally, the frequency of using the Lelylijn reveals a weak positive association (Gamma = 0,078, p = 0,003), indicating that respondents who use the Lelylijn more often are more likely to feel that they learned about the government's choices.

In the Netherlands we should use this method more often to involve residents in government policy

There is a weak negative association between age and the belief that this method should be used more often to involve residents in government policy (Gamma = -0,085, p = 0,017), indicating that older respondents are slightly less likely to support using this method more frequently compared to younger respondents. Gender shows a weak positive association (Gamma = 0,193, p < 0,001), suggesting that women are more likely to support the use of this method compared to men.

Education reveals a weak negative association (Gamma = -0,126, p < 0,001), indicating that respondents with higher levels of education are less likely to support using this method more often compared to those with lower education levels. The frequency of using trains has a weak positive association with this belief (Gamma = 0,095, p = 0,011), suggesting that those who use trains more frequently are more likely to support the use of this method. Additionally, the frequency of using the Lelylijn reveals a weak positive association (Gamma = 0,096, p = 0,045), indicating that respondents who use the Lelylijn

more often are more likely to support using this method more frequently.

Table 4.6: Lelylijn Project Consultation's Cross-tabulation Analysis Results: Acceptance Dimension

Face Validity Dimension		Acceptance								
Statement		In the Netherlands we should use this method more often to involve residents in government policy			If the government involves residents in choices on a large scale through this consultation, then the final decisions about the Lely Line will be more acceptable to me			If the government allows residents to think about these types of choices more often, I will have more confidence in the government's decisions		
Response		Disagree	Neutral	Agree	Disagree	Neutral	Agree	Disagree	Neutral	Agree
Gender	Man	2%	6%	42%	6%	10%	34%	4%	12%	34%
	Woman	1%	5%	44%	6%	12%	32%	4%	14%	31%
	Gamma	0,193			-0,096			-0,088		
	Sig.	<0,001			<0,001			0,002		
Age	18-34 years	2%	4%	40%	6%	9%	32%	4%	10%	32%
	35-64 years	1%	4%	29%	4%	8%	22%	3%	9%	21%
	65 years plus	1%	3%	17%	2%	6%	12%	1%	7%	12%
	Gamma	-0,085			-0,080			-0,145		
	Sig.	0,017			<0,001			<0,001		
Education	Low	1%	2%	22%	4%	4%	17%	2%	6%	17%
	Medium	1%	3%	33%	3%	9%	25%	2%	9%	26%
	High	1%	5%	31%	5%	10%	24%	5%	11%	22%
	Gamma	-0,126			-0,089			-0,155		
	Sig.	<0,001			<0,001			<0,001		
Occupation	Not working	1%	5%	35%	5%	9%	28%	3%	11%	27%
	Part-time	0%	2%	19%	2%	5%	13%	2%	5%	14%
	Full-time	1%	4%	32%	4%	8%	25%	4%	10%	24%
	Gamma	-0,001			-0,011			-0,051		
	Sig.	0,983			0,633			0,036		
Residency	Tenant	1%	4%	22%	4%	5%	18%	2%	6%	18%
	Lodger	0%	1%	12%	1%	3%	9%	0%	3%	10%
	Homeowner	2%	6%	52%	7%	15%	38%	6%	17%	37%
	Gamma	0,055			-0,043			-0,109		
	Sig.	0,163			0,112			<0,001		
Proximity	Not proximate	2%	5%	45%	5%	13%	33%	4%	14%	33%
	Proximate	2%	6%	41%	7%	9%	32%	4%	12%	32%
	Gamma	-0,053			0,060			0,019		
	Sig.	0,205			0,832			0,511		
Frequency to use Train	Rarely	2%	7%	45%	8%	12%	34%	5%	16%	33%
	Occassionally	1%	2%	20%	2%	5%	16%	2%	5%	16%
	(Almost) Daily	33	106	896	112	243	680	65	238	732
	Gamma	0,095			0,090			0,164		
	Sig.	0,011			<0,001			<0,001		
Frequency to use Lelylijn	Rarely	2%	6%	44%	7%	12%	34%	5%	15%	33%
	Occassionally	1%	3%	24%	3%	7%	18%	2%	7%	19%
	(Almost) Daily	20	73	756	65	161	622	57	189	602
	Gamma	0,096			0,099			0,120		
	Sig.	0,045			<0,001			<0,001		
Financial	Not enough	0%	1%	5%	1%	1%	4%	1%	1%	4%
	Enough	2%	8%	60%	8%	15%	46%	5%	18%	46%
	More than Enough	1%	2%	20%	3%	6%	15%	2%	7%	15%
	Gamma	-0,061			-0,067			-0,073		
	Sig.	0,161			0,023			0,016		

If the government involves residents in choices on a large scale through this consultation, then the final decisions about the Lely Line will be more acceptable to me

There is a weak negative association between age and the belief that involving residents on a large scale will make final decisions more acceptable (Gamma = -0,080, $p < 0,001$), suggesting that older respondents are slightly less likely to agree with this statement compared to younger respondents.

Gender shows a weak negative association (Gamma = -0,096, p < 0,001), indicating that men are more likely to increase their acceptance on the final decisions of Lelylijn.

Education has a weak negative association (Gamma = -0,089, p < 0,001), implying that respondents with higher levels of education are less likely to increase their acceptance on the final decisions of Lelylijn compared to those with lower education levels. Financial status also exhibits a weak negative association (Gamma = -0,067, p = 0,023), indicating that those with more financial resources are less likely to increase their acceptance on the final decisions of Lelylijn compared to those with fewer financial resources.

The frequency of using trains shows a weak positive association with this belief (Gamma = 0,090, p < 0,001), suggesting that respondents who use trains more frequently are more likely to increase their acceptance on the final decisions of Lelylijn. Similarly, the frequency of using the Lelylijn reveals a weak positive association (Gamma = 0,099, p < 0,001), indicating that those who use the Lelylijn more often are more likely to increase their acceptance on the final decisions of Lelylijn.

If the government allows residents to think about these types of choices more often, I will have more confidence in the government's decisions

There is a weak negative association between age and the belief that involving residents in decision-making will increase trust in the government's decisions (Gamma = -0,145, p < 0,001), suggesting that older respondents are slightly less likely to agree with this statement compared to younger respondents. Gender also shows a weak negative association (Gamma = -0,088, p = 0,002), indicating that men are more likely to have increase their trust in the government's decisions than women when residents are allowed to participate more frequently in such choices.

Education exhibits a weak negative association (Gamma = -0,155, p < 0,001), implying that respondents with higher levels of education are less likely to have increased confidence in the government's decisions compared to those with lower education levels. Financial status shows a weak negative association (Gamma = -0,073, p = 0,016), indicating that those with more financial resources are less likely to have increased confidence in the government's decisions compared to those with fewer financial resources.

Occupation also has a weak negative association (Gamma = -0,051, p = 0,036), suggesting that those who are employed full-time are less likely to have increased trust in the government's decisions compared to those who are not working. Residency shows a weak negative association (Gamma = -0,109, p < 0,001), indicating that homeowners are less likely to have increased trust in the government's decisions compared to home-renters.

The frequency of using trains displays a weak positive association with this belief (Gamma = 0,164, p < 0,001), suggesting that respondents who use trains more frequently are more likely to have increased trust in the government's decisions. Similarly, the frequency of using the Lelylijn reveals a weak positive association (Gamma = 0,120, p < 0,001), indicating that those who use the Lelylijn more often are more likely to have increased trust in the government's when residents are allowed to think about such choices more often.

4.3.3 Oude Lijn Consultation

Similar crosstabulation analysis were conducted for each respondents' characteristic and the seven face validity statements evaluation in the Oude Lijn consultation. The results for the dimensions of transparency, feasibility, and relevance are presented in Table 4.7, while the results for the acceptance dimension are presented in Table 4.8.

I trust that this research project is conducted in a fair way

There is a weak negative association between age and trust in the project's fairness (Gamma = -0,106, p = 0,005), suggesting that older respondents are slightly less likely to trust the research project. Additionally, there is a weak positive association between residency and trust in the project's fairness (Gamma = 0,117, p = 0,004), indicating that respondents living in owner-occupied homes are slightly more likely to trust the project compared to those in rental housing.

A weak positive association is observed for proximity (Gamma = 0,100, p = 0,031), suggesting that respondents from South Netherlands (close to the project) are slightly more likely to trust the project

compared to those from the rest of the Netherlands. There is also a weak negative association between the frequency of using trains and trust in the project's fairness (Gamma = -0,200, p < 0,001), indicating that respondents who use trains more often are slightly more likely to trust the project.

A weak negative association is found between the frequency of using Oude Lijn and trust in the project's fairness (Gamma = -0,104, p = 0,031), suggesting that respondents who use Oude Lijn more often are slightly more likely to trust the project. Lastly, a weak negative association is observed for getting in/out at any of the mentioned stations (Gamma = -0,138, p = 0,004), indicating that respondents who get in/out at these stations more often are slightly more likely to trust the project.

Table 4.7: Oude Lijn Project Consultation's Cross-tabulation Analysis Results: Transparency, Feasibility, Relevance Dimension

Face Validity Dimension		Transparency			Feasibility			Relevance					
Statement		I trust this is an honest investigation			I found the consultation difficult to understand			I thought it was an important topic to give my opinion			By participating in this consultation, I have learned about the choices that the government has to make on this subject		
Response		Disagree	Neutral	Agree	Disagree	Neutral	Agree	Disagree	Neutral	Agree	Disagree	Neutral	Agree
Gender	Man	1%	5%	42%	36%	8%	5%	2%	9%	38%	3%	12%	34%
	Woman	1%	6%	44%	37%	10%	4%	3%	12%	36%	2%	12%	37%
	Gamma	0,008			0,017			-0,159			0,085		
	Sig.	0,869			0,620			<,001			0,011		
Age	18-34 years	1%	4%	41%	34%	8%	4%	3%	8%	35%	2%	10%	34%
	35-64 years	1%	4%	28%	24%	5%	3%	2%	7%	24%	2%	8%	23%
	65 years or older	0%	3%	17%	15%	4%	1%	1%	5%	14%	1%	6%	13%
	Gamma	-0,106			0,004			-0,056			-0,108		
	Sig.	0,005			0,901			0,051			<,001		
Education	Low	0%	3%	21%	15%	6%	3%	2%	8%	16%	1%	7%	17%
	Medium	1%	4%	33%	27%	7%	3%	3%	8%	27%	1%	9%	27%
	High	1%	4%	33%	31%	4%	3%	1%	5%	31%	3%	8%	26%
	Gamma	0,073			-0,316			0,306			0,015		
	Sig.	0,060			<,001			0,000			0,584		
Occupation	Not working	1%	5%	35%	28%	9%	4%	2%	10%	28%	2%	11%	28%
	Part-time	0%	2%	18%	15%	4%	2%	1%	5%	15%	1%	5%	15%
	Full-time	1%	4%	34%	30%	5%	3%	2%	6%	31%	2%	9%	27%
	Gamma	0,055			-0,150			0,176			0,024		
	Sig.	0,161			<,001			<,001			0,390		
Residency	Rental private	0%	1%	12%	10%	2%	1%	1%	3%	10%	1%	3%	10%
	Rental social	1%	4%	22%	17%	6%	3%	2%	7%	18%	1%	7%	18%
	Owner occupied	1%	6%	53%	46%	10%	4%	3%	12%	45%	3%	14%	42%
	Gamma	0,117			-0,142			0,113			-0,011		
	Sig.	0,004			<,001			<,001			0,708		
Proximity	Rest of Netherlands	1%	6%	39%	31%	10%	5%	4%	13%	29%	2%	12%	32%
	South Netherlands	1%	5%	48%	43%	8%	4%	1%	8%	45%	3%	13%	38%
	Gamma	0,100			-0,272			0,467			-0,034		
	Sig.	0,031			<,001			0,000			0,308		
Frequency of Using Train	At least a few times a month	1%	4%	45%	39%	7%	5%	2%	7%	42%	3%	10%	37%
	Less often or never	1%	7%	42%	34%	11%	4%	4%	14%	32%	2%	14%	33%
	Gamma	-0,200			0,180			-0,446			-0,121		
	Sig.	<,001			<,001			<,001			<,001		
Frequency of Using Oude Lijn	At least a few times a month	1%	3%	32%	29%	4%	4%	1%	4%	32%	3%	7%	27%
	Less often or never	1%	8%	54%	45%	13%	5%	5%	17%	41%	3%	17%	44%
	Gamma	-0,104			0,144			-0,553			-0,066		
	Sig.	0,031			<,001			<,001			0,058		
Live/work at any of Mentioned Stations*	Live/work at any station	1%	4%	32%	27%	5%	4%	1%	5%	31%	3%	8%	26%
	Never Live/work at any station	1%	8%	54%	46%	12%	5%	4%	16%	43%	3%	16%	45%
	Gamma	-0,065			-0,02			-0,423			0,012		
	Sig.	0,179			0,569			<,001			0,722		
Get in/out at any of Mentioned Stations*	Get in/out at any	1%	3%	33%	27%	5%	5%	1%	4%	32%	2%	8%	27%
	Never get in/out at any	1%	8%	54%	46%	12%	4%	4%	17%	42%	3%	16%	44%
	Gamma	-0,138			-0,023			-0,499			-0,043		
	Sig.	0,004			0,512			<,001			0,221		
	Urban	1%	7%	61%	54%	10%	6%	3%	12%	55%	4%	16%	51%

I found the consultation difficult to understand

There is a significant negative association between education level and finding the consultation difficult to understand (Gamma = -0,316, p < 0,001), indicating that respondents with higher education levels are less likely to find the consultation difficult to understand. Additionally, a significant negative

association is observed between occupation and understanding the consultation (Gamma = -0,150, p < 0,001), suggesting that respondents who work full-time are less likely to find the consultation difficult to understand compared to those not working.

A weak negative association exists between residency and finding the consultation difficult to understand (Gamma = -0,142, p < 0,001). This suggests that respondents living in owner-occupied homes are less likely to find the consultation difficult to understand compared to those in rental housing. Proximity also shows a moderate negative association with understanding the consultation (Gamma = -0,272, p < 0,001), indicating that respondents from South Netherlands are less likely to find the consultation difficult to understand compared to those from the rest of the Netherlands.

There is a significant positive association between the frequency of using trains and finding the consultation difficult to understand (Gamma = 0,180, p < 0,001), suggesting that respondents who use trains less often are more likely to find the consultation difficult to understand. A significant positive association is also found for the frequency of using Oude Lijn (Gamma = 0,144, p < 0,001), indicating that respondents who use Oude Lijn less often are more likely to find the consultation difficult to understand. Lastly, a significant positive association is observed for area of living (Gamma = 0,257, p < 0,001), suggesting that respondents living in rural areas are more likely to find the consultation difficult to understand compared to those in urban areas.

I thought it was an important topic to give my opinion on

There is a significant negative association between gender and finding the topic important (Gamma = -0,159, p < 0,001). This suggests that women are slightly less likely to find the topic important compared to men. Additionally, There is a strong positive association between education level and finding the topic important (Gamma = 0,306, p = 0,000), indicating that respondents with higher education levels are more likely to find the topic important.

Occupation also shows a significant positive association with finding the topic important (Gamma = 0,176, p < 0,001), suggesting that respondents who are employed full-time are more likely to find the topic important compared to those who are not working. Similarly, residency exhibits a weak positive association with finding the topic important (Gamma = 0,113, p < 0,001), indicating that respondents who own their homes are more likely to find the topic important compared to those in rental housing.

Proximity shows a moderate positive association with finding the topic important (Gamma = 0,467, p < 0,001). This suggests that respondents from South Netherlands are more likely to find the topic important compared to those from the rest of the Netherlands. There is a moderate negative association between the frequency of using trains and finding the topic important (Gamma = -0,446, p < 0,001), indicating that respondents who use trains less often are less likely to find the topic important. Similarly, the frequency of using Oude Lijn also shows a strong negative association with finding the topic important (Gamma = -0,553, p < 0,001), suggesting that respondents who use Oude Lijn less often are less likely to find the topic important.

Living or working at any of the mentioned stations has a moderate negative association with finding the topic important (Gamma = -0,423, p < 0,001) with higher level being never live or work near any of the stations. This indicates that respondents who never live or work at these stations are less likely to find the topic important. Similarly, getting in or out at any of the mentioned stations shows a moderate negative association with finding the topic important (Gamma = -0,499, p < 0,001), suggesting that respondents who get in or out at these stations are more likely to find the topic important. Lastly, there is a moderate negative association between the area of living and finding the topic important (Gamma = -0,367, p < 0,001), indicating that respondents living in urban areas are more likely to find the topic important compared to those living in rural areas.

By participating in this consultation, I learned about the choices the Province must make on this topic

There is a weak positive association between gender and learning about the government's choices (Gamma = 0,085, p = 0,011), suggesting that women are slightly more likely to feel that they have learned about the government's choices compared to men. A weak negative association is observed between age and learning about the government's choices (Gamma = -0,108, p < 0,001), indicating that older respondents are slightly less likely to feel that they have learned about the government's choices compared to younger respondents.

The frequency of using trains shows a weak negative association with learning about the government's choices (Gamma = -0,121, p < 0,001), suggesting that respondents who use trains less often are slightly less likely to feel that they have learned about the government's choices compared to those who use trains more frequently. Lastly, there is a weak negative association between the area of living and learning about the government's choices (Gamma = -0,118, p = 0,001), suggesting that respondents living in rural areas are slightly less likely to feel that they have learned about the government's choices compared to those living in urban areas.

Table 4.8: Oude Lijn Project Consultation's Cross-tabulation Analysis Results: Acceptance Dimension

Face Validity Dimension		Acceptance								
Statement		This method should be used more often to involve residents in public policy			If many people participate in this consultation, then the final decisions on this subject are better accepted for me			If the government often allows residents to think along about these types of choices more often, then I get more confidence in the decisions of the government		
Response		Disagree	Neutral	Agree	Disagree	Neutral	Agree	Disagree	Neutral	Agree
Gender	Man	2%	6%	41%	4%	12%	33%	3%	11%	35%
	Woman	1%	7%	43%	3%	14%	35%	3%	11%	37%
	Gamma	-0,017			0,021			0,045		
	Sig.	0,687			0,514			0,181		
Age	18-34 years	2%	5%	39%	3%	11%	32%	3%	10%	34%
	35-64 years	1%	5%	28%	2%	9%	22%	2%	8%	23%
	65 years or older	0%	3%	17%	1%	6%	13%	1%	4%	15%
	Gamma	0,012			-0,032			-0,023		
	Sig.	0,737			0,240			0,411		
Education	Low	1%	3%	21%	1%	7%	17%	2%	5%	18%
	Medium	1%	5%	31%	2%	9%	26%	1%	8%	28%
	High	1%	4%	32%	3%	10%	25%	3%	9%	26%
	Gamma	0,097			-0,035			-0,026		
	Sig.	0,006			0,200			0,375		
Occupation	Not working	1%	6%	34%	2%	11%	27%	2%	8%	30%
	Part-time	1%	3%	18%	1%	6%	14%	1%	5%	15%
	Full-time	1%	4%	33%	3%	9%	27%	3%	9%	27%
	Gamma	0,085			0,033			-0,047		
	Sig.	0,016			0,219			0,096		
Residency	Rental private	1%	2%	12%	1%	3%	9%	1%	3%	10%
	Rental social	1%	4%	21%	1%	7%	18%	2%	6%	19%
	Owner occupied	1%	7%	51%	4%	15%	41%	4%	13%	43%
	Gamma	0,092			0,016			-0,010		
	Sig.	0,018			0,585			0,746		
Proximity	Rest of Netherlands	2%	6%	37%	3%	12%	31%	3%	10%	33%
	South Netherlands	1%	6%	47%	4%	13%	38%	3%	12%	39%
	Gamma	0,165			0,033			-0,003		
	Sig.	<,001			0,313			0,940		
Frequency of Using Train	At least a few times a month	1%	6%	43%	3%	12%	35%	3%	11%	37%
	Less often or never	2%	7%	41%	3%	13%	33%	3%	11%	35%
	Gamma	-0,141			-0,036			-0,065		
	Sig.	<,001			0,262			0,054		
Frequency of Using Oude Lijn	At least a few times a month	1%	4%	32%	2%	9%	26%	2%	8%	27%
	Less often or never	2%	9%	52%	4%	17%	42%	4%	14%	45%
	Gamma	-0,138			-0,066			-0,031		
	Sig.	0,002			0,051			0,374		
*Live/work at any of Mentioned Stations**	Live/work at any station	1%	4%	31%	3%	9%	25%	2%	8%	26%
	Never Live/work at any station	2%	8%	53%	4%	17%	43%	4%	14%	45%
	Gamma	-0,025			-0,015			0,008		
	Sig.	0,574			0,667			0,809		
*Get in/out at any of Mentioned Stations**	Get in/out at any	1%	4%	32%	3%	9%	26%	2%	8%	27%
	Never get in/out at any	2%	9%	52%	4%	17%	43%	4%	14%	45%
	Gamma	-0,129			-0,027			-0,035		
	Sig.	0,004			0,430			0,324		
Area of Living	Urban	1%	9%	60%	4%	18%	48%	4%	16%	50%
	Rural	2%	4%	24%	2%	8%	20%	2%	6%	21%
	Gamma	-0,182			-0,075			-0,016		
	Sig.	<,001			0,036			0,658		

This method should be used more often to involve residents in government policies

There is a significant weak positive association between education level and the belief that this method should be used more often to involve residents in public policy (Gamma = 0.097, p = 0.006). This suggests that respondents with higher education levels are slightly more likely to support the frequent

use of this method. Additionally, a weak positive association is observed between occupation and this belief (Gamma = 0.085, p = 0.016), indicating that respondents who are working full time are slightly more likely to believe this method should be used more often compared to those not working.

Residency also shows a weak positive association with the belief that this method should be used more often (Gamma = 0.092, p = 0.018), suggesting that owner-occupied residents are slightly more likely to support the frequent use of this method compared to those in private or social rental housing. Similarly, proximity reveals a weak positive association with this belief (Gamma = 0.165, p < 0.001), indicating that respondents from the South Netherlands are more likely to support the frequent use of this method compared to those from the rest of the Netherlands.

There is a weak negative association between the frequency of using trains (Gamma = -0.141, p < 0.001) and the frequency of using Oude Lijn (Gamma = -0.138, p = 0.002) with the belief in the method's frequent use. This suggests that respondents who use trains and Oude Lijn less often are less likely to support the frequent use of this method. Furthermore, there is a weak negative association for getting in or out at any of the mentioned stations and the belief in the method's frequent use (Gamma = -0.129, p = 0.004). This suggests that respondents who never get in or out at these stations are less likely to support the frequent use of this method. Lastly, the area of living shows a weak negative association with this belief (Gamma = -0.182, p < 0.001), indicating that respondents living in rural areas are less likely to believe this method should be used more often compared to those living in urban areas.

If many people participate in this consultation, the final decisions on this topic are more acceptable to me

There is a weak negative association between the area of living and the belief that final decisions are better accepted if many people participate in the consultation (Gamma = -0.075, p = 0.036). This suggests that respondents living in rural areas are slightly less likely to increase their acceptance on the final decisions compared to those living in urban areas. Other variables did not show significant associations, meaning meaningful association between them and the increase on the final decisions cannot be concluded.

If the government involves residents more often in thinking about these kinds of choices, I will gain more confidence in the government's decisions

The analysis of the statement "If the government often allows residents to think along about these types of choices more often, then I get more confidence in the decisions of the government" reveals that none of the demographic or behavioral attributes show significant associations with the responses. The lack of significant associations across all examined attributes suggests that respondents' increase of trust in government decisions involving resident input is broadly consistent across different demographic and behavioral groups.

4.3.4 Mobility Vision Consultation

Similar crosstabulation analysis were conducted for each respondents' characteristic and the seven face validity statements evaluations in the Mobility Vision consultation. The results for the dimensions of transparency, feasibility, and relevance are presented in Table 4.9, while the results for the acceptance dimension are presented in Table 4.10.

I trust that this research project is conducted in a fair way

The analysis reveals a weak positive association between age and trust in the research project's fairness (Gamma = 0,083, p = 0,017). This indicates that older respondents in this consultation are slightly more likely to trust that the research project is conducted in a fair way. Additionally, there is a weak positive association between education level and trust in the project's fairness (Gamma = 0,122, p = 0,000), suggesting that those with higher education levels have greater trust in the project's fairness. Conversely, there is a weak negative association between facing accessibility problems and trust in the research project's fairness (Gamma = -0,152, p = 0,000), indicating that respondents facing more accessibility problems are slightly less likely to trust the research's fairness. Other variables did not show significant associations, meaning we cannot conclude a meaningful association between them and trust in the project's fairness.

I found the consultation difficult to understand

The analysis shows a moderate negative association between age and finding the consultation difficult

to understand (Gamma = -0,220, p = 0,000). This indicates that older respondents in this consultation are less likely to find the consultation difficult to understand. There is also a moderate negative association between education level and finding the consultation difficult to understand (Gamma = -0,355, p = 0,000), suggesting that higher education levels decrease the likelihood of finding the consultation difficult to understand. Additionally, there is a weak positive association between occupation and finding the consultation difficult to understand (Gamma = 0,075, p = 0,023), with full-time workers slightly more likely to find the consultation difficult compared to those not working. Furthermore, a weak negative association is observed between area of living and finding the consultation difficult to understand (Gamma = -0,069, p = 0,038), indicating that rural respondents are slightly less likely to find the consultation difficult compared to urban respondents. Other variables did not show significant associations, meaning we cannot conclude a meaningful association between them and finding the consultation difficult to understand.

Table 4.9: Mobility Vision Project Consultation's Cross-tabulation Analysis Results: Transparency, Feasibility and Relevance Dimension

Face Validity Dimension		Transparency			Feasibility			Relevance					
Statement		I trust that this research project is conducted in a fair way			I found the consultation difficult to understand			I thought it was an important topic to give my opinion on			By participating in this consultation, I learned about the choices the Province must make on this topic		
Response		Disagree	Neutral	Agree	Disagree	Neutral	Agree	Disagree	Neutral	Agree	Disagree	Neutral	Agree
Gender	Man	2%	6%	42%	37%	7%	6%	1%	5%	43%	5%	14%	30%
	Woman	1%	7%	43%	36%	11%	4%	1%	8%	42%	4%	15%	32%
	Gamma	-0,011			0,061			-0,207			0,062		
	Sig.	0,796			0,057			0,000			0,032		
Age	18-34 years	2%	6%	39%	31%	9%	6%	2%	7%	38%	4%	13%	29%
	35-64 years	1%	5%	28%	26%	5%	2%	1%	3%	29%	3%	10%	20%
	65 years or older	0%	2%	18%	16%	4%	1%	0%	2%	18%	1%	7%	13%
	Gamma	0,083			-0,220			0,159			-0,032		
	Sig.	0,017			0,000			0,000			0,181		
Education	Low	1%	4%	21%	15%	7%	4%	1%	4%	20%	1%	7%	17%
	Medium	1%	5%	31%	27%	7%	4%	1%	5%	31%	2%	11%	24%
	High	1%	4%	32%	31%	4%	3%	1%	3%	33%	5%	12%	21%
	Gamma	0,122			-0,355			0,259			-0,176		
	Sig.	0,000			0,000			0,000			0,000		
Occupation	Not working	1%	5%	32%	28%	7%	3%	1%	5%	32%	3%	12%	23%
	Full-time	2%	8%	53%	45%	11%	7%	2%	8%	53%	6%	18%	38%
	Gamma	-0,069			0,075			-0,003			-0,009		
	Sig.	0,109			0,023			0,940			0,761		
Area of Living	Urban	2%	8%	54%	46%	11%	7%	2%	8%	54%	6%	19%	39%
	Rural	1%	5%	31%	27%	7%	3%	1%	5%	31%	3%	11%	23%
	Gamma	0,021			-0,069			0,025			0,044		
	Sig.	0,631			0,038			0,562			0,145		
Frequency of Facing Acessibility Problem	Low	2%	7%	55%	47%	11%	6%	1%	8%	54%	5%	19%	40%
	Moderate	1%	4%	23%	19%	5%	3%	1%	3%	23%	3%	8%	16%
	High	0%	2%	7%	6%	2%	1%	0%	1%	7%	1%	3%	5%
	Gamma	-0,152			0,053			-0,007			-0,075		
	Sig.	0,000			0,082			0,859			0,007		

I thought it was an important topic to give my opinion on

The analysis indicates a moderate negative association between gender and finding the topic important (Gamma = -0,207, p = 0,000). This suggests that women in this consultation are less likely than men to find the topic important. A weak positive association is observed between age and finding the topic important (Gamma = 0,159, p = 0,000), indicating that older respondents are more likely to find the topic important. There is also a moderate positive association between education level and finding the topic important (Gamma = 0,259, p = 0,000), showing that higher education levels are associated with a greater likelihood of finding the topic important. Other variables did not show significant associations, meaning we cannot conclude a meaningful association between them and finding the topic important.

By participating in this consultation, I learned about the choices the Province must make on this topic

The analysis reveals a weak positive association between gender and learning about provincial choices (Gamma = 0,062, p = 0,032). This indicates that women in this consultation are slightly more likely than

men to feel that they learned about the choices the Province must make by participating in the consultation. There is a weak negative association between education level and learning about provincial choices (Gamma = -0,176, p = 0,000), suggesting that higher education levels are associated with a decreased likelihood of feeling that they learned about provincial choices. Similarly, a weak negative association is observed between facing accessibility problems and learning about provincial choices (Gamma = -0,075, p = 0,007), indicating that those facing more accessibility problems are slightly less likely to feel that they learned about the choices the Province must make. Other variables did not show significant associations, meaning we cannot conclude a meaningful association between them and learning about provincial choices.

Table 4.10: Mobility Vision Project Consultation's Cross-tabulation Analysis Results: Acceptance Dimension

Face Validity Dimension		Acceptance								
Statement		This method should be used more often to involve residents in government policies			If many people participate in this consultation, the final decisions on this topic are more acceptable to me			If the government involves residents more often in thinking about these kinds of choices, I will have more confidence in the		
Response		Disagree	Neutral	Agree	Disagree	Neutral	Agree	Disagree	Neutral	Agree
Gender	Man	2%	8%	40%	5%	13%	31%	4%	12%	34%
	Woman	1%	10%	40%	3%	18%	30%	3%	15%	32%
	Gamma	-0,055			-0,037			-0,070		
	Sig.	0,137			0,198			0,020		
Age	18-34 years	1%	9%	36%	4%	13%	28%	3%	12%	30%
	35-64 years	1%	5%	27%	3%	11%	20%	3%	9%	21%
	65 years or older	1%	4%	17%	1%	7%	13%	1%	6%	14%
	Gamma	0,062			0,010			-0,008		
	Sig.	0,049			0,683			0,751		
Education	Low	1%	6%	19%	1%	8%	17%	1%	7%	18%
	Medium	1%	6%	30%	3%	12%	23%	3%	9%	25%
	High	1%	6%	30%	4%	11%	22%	4%	10%	23%
	Gamma	0,085			-0,112			-0,114		
	Sig.	0,007			0,000			0,000		
Occupation	Not working	1%	7%	30%	2%	12%	23%	2%	10%	25%
	Full-time	2%	11%	49%	6%	18%	38%	5%	16%	41%
	Gamma	-0,034			-0,029			-0,041		
	Sig.	0,372			0,339			0,186		
Area of Living	Urban	2%	11%	50%	6%	19%	38%	5%	17%	41%
	Rural	1%	6%	29%	2%	11%	23%	2%	10%	24%
	Gamma	0,061			0,063			0,055		
	Sig.	0,110			0,038			0,080		
Frequency of Facing Acessibility Problem	Low	2%	11%	51%	5%	20%	40%	5%	17%	42%
	Moderate	1%	5%	21%	2%	9%	16%	2%	7%	18%
	High	0%	1%	7%	1%	2%	5%	0%	2%	6%
	Gamma	-0,032			-0,051			0,010		
	Sig.	0,373			0,069			0,737		

This method should be used more often to involve residents in government policies

The analysis reveals a weak positive association between age and the belief that this method should be used more often to involve residents in government policies (Gamma = 0,062, p = 0,049). This indicates that older respondents in this consultation are slightly more likely to believe that this method should be used more often. Additionally, there is a weak positive association between education level and this belief (Gamma = 0,085, p = 0,007), suggesting that those with higher education levels are slightly more likely to believe that this method should be used more often. Other variables did not show significant associations, meaning we cannot conclude a meaningful association between them and the belief that this method should be used more often to involve residents in government policies.

If many people participate in this consultation, the final decisions on this topic are more acceptable to me

The analysis indicates a weak negative association between education level and the belief that the final decisions are more acceptable with more participation (Gamma = -0,112, p = 0,000), suggesting that higher education levels are associated with a decreased likelihood increasing their acceptance on the final decisions. Additionally, there is a weak positive association between area of living and this belief (Gamma = 0,063, p = 0,038), indicating that rural respondents are slightly more likely to increase their acceptance on the final decisions compared to urban respondents. Other variables did not show significant associations.

If the government involves residents more often in thinking about these kinds of choices, I will have more confidence in the government's decisions

The analysis reveals a weak negative association between gender and the belief that more involvement of residents increases trust in government (Gamma = -0,070, p = 0,020). This indicates that women in this consultation are slightly less likely than men to experience that participating in PVE increases their trust in the government. Similarly, a weak negative association is observed between education level and this belief (Gamma = -0,114, p = 0,000), suggesting that higher education levels are associated with a less likely experience that participating in PVE increases their trust in the government. Other variables did not show significant associations, meaning we cannot conclude a meaningful association between them and the belief that more involvement of residents increases trust in government decisions.

4.3.5 Comparison of Crosstabulation Analysis Across Consultations

The individual crosstabulation analysis for each PVE consultation reveals numerous significant associations between the characteristics of the respondents and their assessment of the face validity statements in each consultation. However, there are instances where this strong correlation is consistent throughout consultations and instances where it is not. The comparisons of these associations across three PVE consultations are as follows:

1. **Age**
   Negative: There is a significant consistent evaluation that as someone becomes older, they tend to think that this consultation is steering.
   Positive: There are considerable positive evaluations that indicate that older individuals are more inclined to embrace and have confidence in government decisions, despite the fact that consistency cannot be verified. Additionally, older individuals are less inclined to perceive the consultation as difficult.
   Inconsistency: Age appears to associate differently with the face validity statements about the consultation being fair, important, and be used more often.

2. **Gender**
   Negative: Men are less likely to think they learned about the government's choices through this consultation. They also tend to think this consultation is quite steering.
   Positive: Men tend to think that their participation in PVE will increase their acceptance on the final decisions and their trust in the government.
   Inconsistency: The way in which gender is associated with the face validity statements regarding the topic's importance and the frequency with which this consultation should be used varies depending on the context.

3. **Education**
   Negative: The higher the education level someone has, the less likely they are to increase their acceptance on the final decisions on the project and to increase their trust in the government.
   Positive: People with higher education level tend to think the consultation is easier and the topic is more likely to be important
   Inconsistency: It's not clear how someone's level of education affects their opinion on whether PVE consultation is fair, whether they learned about the government's decisions, and whether this method should be used more.

4. **Occupation**
   Negative: Although the consistency's cannot be checked, full time workers apparently are less

likely to increase their trust in the government.
Positive: Even if the consistency cannot be verified, full-time employees frequently believe that the consultation is not steering.
Inconsistency: Despite many significant association between occupation and the face validity evaluation, there is no consistent pattern whether full time workers evaluate the face validity of PVE more positively or not. The outcome is contingent upon the specific circumstances.

5. **Residency**
   Negative: Even though it's not always the case, people who live in their own homes are less likely to think they learned something from PVE consultations and have less trust in the government's final choices.
   Positive: Homeonwer tend to think that PVE should be used more often.
   Inconsistency: There is no consistent association between the type of house an individual resides in and their assessment of the fairness and significance of this consultation.

6. **Proximity**
   Negative: The closer someone lives to the project, the more likely they think the consultation is steering.
   Positive: There is a constant association between proximity to a project and the likelihood of perceiving it as a fair examination and an important topic. In addition, despite the inability to verify its consistency, someone who lives closer also believe that this method should be employed more often.
   Inconsistency: none.

7. **Frequency of Train Use**
   Negative: none.
   Positive: There is a consistent positive correlation between the frequency of train usage and individuals' perceptions of the fairness, importance, learned about the government's choices, and potential for increased use of PVE consultation. Additionally, despite the absence of consistent evidence, people who travel more often with train are also less likely to think that the consultation is steering and more likely to increase their acceptance and trust on the government's decisions.
   Inconsistency: none.

8. **Frequency of (planned) Line Use**
   Negative: none.
   Positive: There is a consistent positive correlation between the individuals' perceptions of the fairness, importance, and potential for increased use of PVE consultation. Additionally, despite the absence of consistent evidence, the more someone (plan to) use the Line, the more positive they are about learning, being confident and increasing acceptance on the government's decisions.
   Inconsistency: none.

9. **Live/ work near stations**
   Negative: none.
   Positive: Although the consistency across consultations cannot be verified, individuals residing or employed in close proximity to any of the studied stations are less inclined to consider it an important topic.
   Inconsistency: none.

10. **Get in/out at stations**
    Negative: none.
    Positive: While the consistency cannot be verified, people who ever get in/out at any of aforementioned stations are more likely to trust this consultation and think that this is an important topic, and that this approach should be employed more frequently.

Inconsistency: none.

11. **Area of Living**
Negative: none.
Positive: Although the consistency cannot be confirmed, urban residents generally favour for a more regular adoption of this method.
Inconsistency: There is a lack of consistent association between an individual's residential area (urban or rural) and their perception of the ease of consultation, the importance of the topic, their awareness of government choices, and their increase of acceptance of government decisions within the dataset.

12. **Financial**
Negative: Although consistency cannot be checked, people who have more money are less likely to increase their acceptance and trust on the government's decisions.
Positive: Someone who has more than enough money every month are less likely to think the consultation is steering.
Inconsistency: none.

13. **Frequency of Facing Accessibility Problems**
Negative: While it is not possible to verify the consistency, there is a notable association between experiencing accessibility issues and having a less positive perception of the fairness of this consultation and the extent to which one learns from it.
Positive: none.
Inconsistency: none.

4.4 Latent Class Cluster Analysis Results (LCCA)

Latent Class Cluster Analysis (LCCA) was performed to identify groups in each consultation with certain characteristics that collectively evaluate PVE consultations more positively or negatively. The subsequent parts provide a comprehensive analysis of the LCCA outcomes for each PVE consultation.

4.4.1 Lelylijn Consultation

In the Lelylijn consultation dataset, a Latent Class Cluster Analysis (LCCA) was conducted for models ranging from 1-cluster to 10-clusters, resulting in the output represented in Table 4.11. To determine the optimal model, several statistical indicators were employed as depicted in section 2.2.5.

Table 4.11: Output of Latent Class Cluster Analysis in Lelylijn Consultation

Model	LL	BIC(LL)	AIC(LL)	AIC3(LL)	Npar	L^2	Max. BVR	VLMR	Class. Err.	Entropy R^2	Difference			
											LL	BIC(LL)	AIC(LL)	AIC3(LL)
1-Cluster	-29.820	59.834	59.686	59.709	23	48.454	1.520		-	1,00				
2-Cluster	-27.422	55.189	54.926	54.967	41	43.657	213	4.797	0,06	0,77	2.398	-4.645	-4.761	-4.743
3-Cluster	-27.032	54.561	54.182	54.241	59	42.877	228	780	0,05	0,83	390	-628	-744	-726
4-Cluster	-26.770	54.189	53.694	53.771	77	42.354	107	524	0,13	0,70	262	-372	-488	-470
5-Cluster	-26.436	53.673	53.062	53.157	95	41.686	114		0,12	0,74	334	-516	-632	-614
6-Cluster	-26.130	53.212	52.485	52.598	113	41.073	56	613	0,11	0,77	306	-461	-577	-559
7-Cluster	-25.829	52.763	51.921	52.052	131	40.473	49	600	0,17	0,72	300	-449	-564	-546
8-Cluster	-25.540	52.336	51.378	51.527	149	39.894	31	579	0,17	0,75	289	-427	-543	-525
9-Cluster	-25.295	51.998	50.925	51.092	167	39.405	37	489	0,09	0,84	245	-338	-453	-435
10-Cluster	-25.055	51.670	50.481	50.666	185	38.925	40	480	0,08	0,87	240	-328	-444	-426

The complete output of the LCCA showed a consistent decrease in BIC(LL), AIC(LL), and AIC3(LL) values up to the 10-clusters model. Similarly, the Log-Likelihood value increased gradually up to the 10-clusters model. The Max. BVR value, while fluctuating across the 10 different clusters, reached its minimum in the 8-clusters model. In terms of Entropy R^2, the 10-cluster model had the highest value,

aside from the 1-cluster model which inherently has the highest value.

When evaluating the decrease in BIC(LL), the 3-cluster model (-628) and the 5-cluster model (-516) showed the most significant decreases. Therefore, the options were narrowed down to 3-cluster up to 5-cluster models. A deeper evaluation of the cluster profiles revealed that the 4-cluster model provided a more meaningful differentiation among respondents, particularly for those who negatively evaluated the PVE's face validity. Additionally, the percentages of each cluster are not less than 5%. Furthermore, since the primary focus of this study is eventually to improve the future PVE design, particularly for respondents who negatively evaluated the PVE's face validity, the 4-cluster model was chosen. This model provides a more parsimonious solution and offers clearer, more actionable recommendations for addressing the concerns of negative evaluators. The results of the 4-cluster model LCCA are presented in Figure 4.17, and further detailed in Appendix D.1 and D.2. The following is a more detailed description of each cluster.

Cluster 1: Highly Positive Evaluators

The first cluster, comprising 54% of the respondents, primarily consists of individuals who evaluate PVE's face validity positively. Members of this cluster have a high probability (55%) of disagreeing with the negatively formulated statement "The consultation steered my choices in a certain direction" and are very likely to agree with all positively formulated statements, with probabilities ranging from 88% to 98%. They also tend to rate the PVE consultation the highest, with an average rating of 7,93. In terms of demographics, this cluster includes a balanced gender ratio and almost half of them are middle-aged individuals (48%). They are more likely to have a medium education level (40%) or high education level (32%). Financially, most people in this cluster have enough money every month (71%) or more than enough (24%). Although 36% are not working, a significant portion works full-time (40%). Most individuals in this cluster live in owner-occupied homes (67%) and have an equal likelihood of living near or far from the Lelylijn project. More than a half of the respondents in this cluster rarely use both trains (60%) and plan to travel through Lelylijn (56%).

Cluster 2: Skeptical Positive Evaluators

The second cluster, which constitutes 24% of the respondents, displays a positive with a bit more neutral stance towards PVE's face validity. Members of this cluster are more likely to agree with 4 out of 6 positively formulated statements, although still disagree with the negatively formulated statement "The consultation steered my choices in a certain direction" (49%). They rate the PVE consultation relatively high, with an average of 7,51. Demographically, this cluster is similar to the first cluster which constitute mostly middle-aged individuals (45%). Women (52%) are slightly more prevalent in this cluster, and it has the highest proportion of individuals with high education levels across clusters (45%). This group predominantly has enough money every month (64%). Most individuals in this cluster are either working full-time (36%) or not working (44%). They are also more likely to live in owner-occupied homes (70%) and have an higher likelihood of living near the Lelylijn project (54%). More than half of the respondents in this cluster rarely use trains (64%) and plan to travel through Lelylijn (62%). Essentially, what make this group distinct from the first cluster is that the members of this cluster were not as positive in evaluating the face validity regarding learning, increased acceptance and trusts.

Cluster 3: Critical Evaluators - Potential Users

The third cluster, making up 11% of the respondents, represents a neutral to negative stance. Members of this cluster are more likely (53%) to agree with the negatively formulated statement "The consultation steered my choices in a certain direction", while remaining neutral on 1 of the positively formulated statements. They rate the PVE consultation relatively low, with an average rating of 6,08. This cluster predominantly consists of middle-aged individuals (48%) and older adults (36%), with a higher likelihood of being men (53%). This cluster has the least likelihood of young adults across the clusters. The education level distributions are almost balanced across different levels. Most members have sufficient financial means (78%), and work full-time (39%). Although almost half of them are not working (46%). They are more likely to live in owner-occupied homes (83%) located far from the Lelylijn project (69%). Although this cluster has the highest likelihood of rarely using trains (61%), almost half of the members of this group plan to occassionally or frequently travel through Lelylijn (49%).

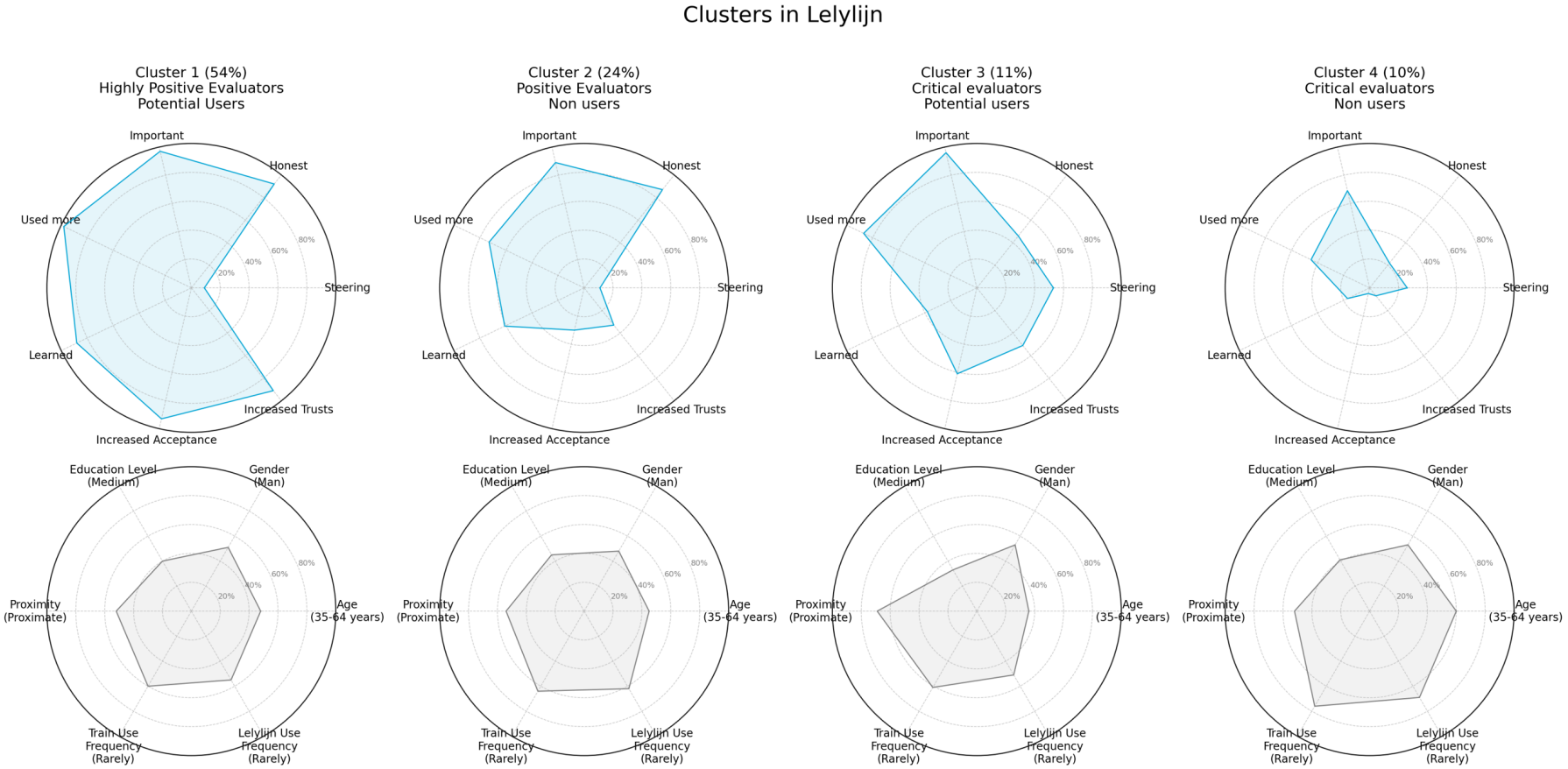


Figure 4.17: Clusters in Lelylijn Dataset

Cluster 4: Critical Evaluators - Non Users

The fourth cluster, which includes 10% of the respondents, evaluates the PVE most negatively. Members of this cluster are likely to disagree with 3 out of 6 positively formulated statements and be neutral with the negatively formulated statement ”The consultation steered my choices in a certain direction” (46%). Similar to the third cluster, they rate the consultation poorly, with the lowest average rating of 5,58. Demographically, this cluster consists mostly of middle-aged individuals (60%) and young adults (24%). Men are slightly overrepresented in this cluster (53%). Most members have a high education level (41%) or medium education level (29%). Similar to other clusters, the majority of the members have enough money every month (69%). Most individuals in this cluster are working full-time (43%) and live far from the Lelylijn project (52%). The members of this cluster are the least likely to use train (76%) and plan to use Lelylijn (69%) compared to other clusters.

4.4.2 Oude Lijn Consultation

In the Oude Lijn consultation dataset, a Latent Class Cluster Analysis (LCCA) was conducted for models ranging from 1-cluster to 10-clusters, resulting in the output represented in Table 4.12. To determine the optimal model, several statistical indicators were employed as depicted in section 2.2.5.

Table 4.12: Output of Latent Class Cluster Analysis in Oude Lijn Consultation

Model	LL	BIC(LL)	AIC(LL)	AIC3(LL)	Npar	L^2	Max. BVR	VLMR	Class. Err.	Entropy R^2	Difference			
											LL	BIC(LL)	AIC(LL)	AIC3(LL)
1-Cluster	-25.012	50.215	50.070	50.093	23	39.119	1.434		-	1,00				
2-Cluster	-22.795	45.947	45.676	45.719	43	34.686	181	4.433	0,05	0,78	2.217	-4.267	-4.393	-4.373
3-Cluster	-22.489	45.502	45.105	45.168	63	34.074	170	611	0,13	0,67	306	-445	-571	-551
4-Cluster	-22.230	45.150	44.626	44.709	83	33.556	57	519	0,15	0,68	259	-353	-479	-459
5-Cluster	-22.043	44.941	44.291	44.394	103	33.181	51	375	0,16	0,70	187	-208	-335	-315
6-Cluster	-21.902	44.827	44.051	44.174	123	32.901	23	281	0,17	0,70	140	-115	-241	-221
7-Cluster	-21.792	44.772	43.870	44.013	143	32.680	25	220	0,18	0,71	110	-54	-180	-160
8-Cluster	-21.704	44.762	43.734	43.897	163	32.504	25	176	0,18	0,70	88	-10	-136	-116
9-Cluster	-21.624	44.768	43.614	43.797	183	32.344	15	160	0,20	0,69	80	6	-120	-100
10-Cluster	-21.558	44.802	43.522	43.725	203	32.212	17	132	0,20	0,70	66	34	-92	-72

The complete output of the LCCA showed a consistent decrease in BIC(LL), AIC(LL), and AIC3(LL) values up to the 8-clusters model, which then started to increase again from the 9-clusters model for BIC(LL) value. Similarly, the Log-Likelihood value increased gradually up to the 10-clusters model. The Max. BVR value, while fluctuating across the 10 different clusters, reached its minimum in the 9-clusters model. In terms of Entropy R^2, the 2-clusters models had the highest value, aside from the 1-cluster model which inherently has the highest value.

When evaluating the decrease in BIC(LL), the 3-cluster model (-445) up to the 5-cluster model (-208) showed the most significant decreases. Therefore, the options were narrowed down to 3-clusters up to 5-cluster models. A deeper evaluation of the cluster profiles revealed that the 4-cluster model provided a more meaningful differentiation among respondents, particularly for those who negatively evaluated the PVE’s face validity. Furthermore, since the primary focus of this study is eventually to improve the future PVE design, particularly for respondents who negatively evaluated the PVE’s face validity, the 4-cluster model was chosen. This model provides a more parsimonious solution and offers clearer, more actionable recommendations for addressing the concerns of negative evaluators. The results of the 4-cluster model LCCA are presented in Figure 4.18, and further detailed in Appendix D.3 and D.4. The following is a more detailed description of each cluster.

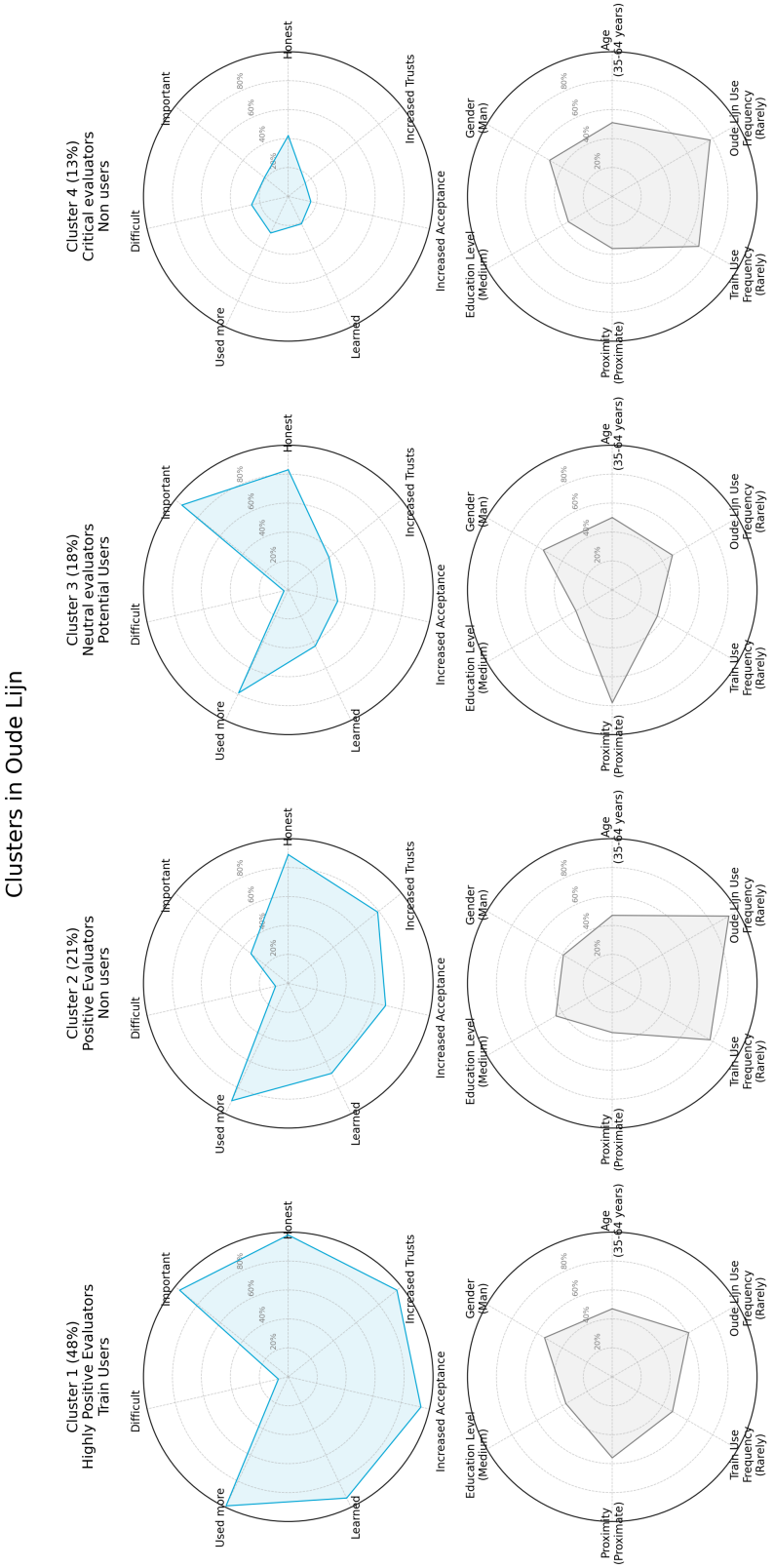


Figure 4.18: Clusters in Oude Lijn Dataset

Cluster 1: Highly Positive Evaluators - Train Users

The first cluster, comprising 48% of the respondents, primarily consists of individuals who evaluate PVE's face validity positively. Members of this cluster have a high probability (77%) of disagreeing with the negatively formulated statement "I found the consultation difficult to understand" and are very likely to agree with all other positively formulated statements, with probabilities ranging from 93% to 99%. They also tend to rate the PVE consultation the highest, with an average rating of 8,19. In terms of demographics, this cluster is slightly dominated by men (54%) and almost half of them are middle-aged individuals (47%). They are more likely to have a high education level (36%) or medium education level (37%). Although 39% are not working, a similar portion works full-time (41%). Most individuals in this cluster live in owner-occupied homes (62%) and are more likely to live in urban (73%) area. More than half of the respondents in this cluster use trains frequently (52%) and but they are less likely to travel through Oude Lijn (61%). They are also less likely to ever get in or out (62%) and live or work near stations in the study (62%).

Cluster 2: Positive Evaluators - Non Users

The second cluster, which constitutes 21% of the respondents, also displays a positive stance towards Oude Lijn PVE's face validity. Members of this cluster are more likely to agree with 5 out of 6 positively formulated statements, although they remain neutral on the statement "I thought it was an important topic to give my opinion" (55%) and disagree with the negatively formulated statement "I found the consultation difficult to understand" (71%). They rate the PVE consultation relatively high, with an average of 7,47. Demographically, this cluster consists mainly of middle-aged individuals (47%), with the least likelihood of young adults across the clusters. Women (61%) are more prevalent in this cluster, and it has a higher proportion of individuals with medium education levels (45%). This group predominantly has enough money every month (64%). Most individuals in this cluster are not working (47%). They are also less likely to live in South Holland (66%) and in rural areas (45%). More than half of the respondents in this cluster rarely use trains (78%) and plan to travel through Oude Lijn less often or never (93%). Consequently, they are less likely to ever get in or out (90%), and live or work near any stations along Oude Lijn (88%). Out of all the clusters, this particular cluster is the most likely not to utilize the line.

Cluster 3: Neutral Evaluators - Potential Users

The third cluster, making up 18% of the respondents, represents a neutral stance. Members of this cluster has the highest probability (85%) to disagree with the negatively formulated statement "this consultation is difficult to understand" across all clusters. They tend to be neutral in 3 out of 6 positively formulated statements and agree with the rest. They rate the PVE consultation slightly higher than the 2nd cluster, with an average rating of 7,50. This cluster predominantly consists of middle-aged individuals (50%) and young adults (27%), with a higher likelihood of being men (55%). More than half of the members have high education level (59%) or medium (29%), aligning with their high probability of disagreeing that this consultation is difficult. Most members work full-time (49%), although almost half of them are not working (31%). They are more likely to live in owner-occupied homes (73%) located close to the project in South Holland (78%). This cluster exhibits the highest likelihood of having more regular train usage (64%), utilization of the Oude Lijn (52%), entering or exiting from the stations (51%), and residing or working in close proximity to the stations (49%), in comparison to the other clusters.

Cluster 4: Critical Evaluators - Non Users

The fourth cluster, which includes 13% of the respondents, evaluates the PVE least positively. Members of this cluster are likely to be neutral with all 6 positively formulated statements, although more likely to disagree with the statement "The consultation steered my choices in a certain direction" (48%). The members tend to rate the consultation poorly, with the lowest average rating of 6,52. Demographically, this cluster consists mostly of middle-aged individuals (51%) and young adults (26%). There is a balanced proportion of gender in this group. This cluster has the the highest probability of being low educated individual, compared to other clusters. This aligns with their least probability to disagree that this consultation is difficult, compared to other clusters. Most individuals in this cluster are not working (47%). The members of this cluster are less likely to use trains (69%), plan to use Oude Lijn (78%), get in or out from the stations (77%), and live or work near the stations (73%). They are also more likely to live far from the Oude Lijn (64%).

4.4.3 Mobility Vision Consultation

In the Mobility Vision consultation dataset, a Latent Class Cluster Analysis (LCCA) was conducted for models ranging from 1-cluster to 10-clusters, resulting in the output represented in Table 4.13. To determine the optimal model, several statistical indicators were employed as depicted in section 2.2.5.

Table 4.13: Output of Latent Class Cluster Analysis in Mobility Vision Consultation

Model	LL	BIC(LL)	AIC(LL)	AIC3(LL)	Npar	L^2	Max. BVR	VLMR	Class Error	Entropy R^2	Difference			
											LL	BIC(LL)	AIC(LL)	AIC3(LL)
1-Cluster	-26.962	54.116	53.970	53.993	23	28.998	1.688		-	1,00				
2-Cluster	-24.385	49.088	48.847	48.885	38	23.845	170	5.153	0,06	0,78	2.577	-5.028	-5.123	-5.108
3-Cluster	-23.952	48.347	48.010	48.063	53	22.978	118	867	0,12	0,71	433	-741	-837	-822
4-Cluster	-23.696	47.961	47.528	47.596	68	22.467	41	512	0,13	0,72	256	-386	-482	-467
5-Cluster	-23.497	47.688	47.161	47.244	83	22.069	42	397	0,12	0,75	199	-272	-367	-352
6-Cluster	-23.377	47.573	46.950	47.048	98	21.828	19	241	0,15	0,73	121	-116	-211	-196
7-Cluster	-23.285	47.514	46.796	46.909	113	21.644	21	184	0,15	0,73	92	-59	-154	-139
8-Cluster	-23.202	47.473	46.659	46.787	128	21.478	18	167	0,19	0,70	83	-41	-137	-122
9-Cluster	-23.132	47.459	46.550	46.693	143	21.339	12	139	0,19	0,71	69	-14	-109	-94
10-Cluster	-23.084	47.488	46.484	46.642	158	21.242	9	97	0,19	0,71	48	29	-67	-52

When evaluating the decrease in BIC(LL), the 3-cluster model (-741) up to the 5-cluster model (-272) showed the most significant decreases. Therefore, the options were narrowed down to 3-cluster up to 5-cluster models. A deeper evaluation of the cluster profiles revealed that the 4-cluster model provided a more meaningful differentiation among respondents, particularly for those who negatively evaluated the PVE's face validity. Additionally, the percentages of each cluster are not less than 5%. Furthermore, since the primary focus of this study is eventually to improve the future PVE design, particularly for respondents who negatively evaluated the PVE's face validity, the 4-cluster model was chosen. This model provides a more parsimonious solution and offers clearer, more actionable recommendations for addressing the concerns of negative evaluators. The results of the 4-cluster model LCCA are presented in Figure 4.19, and further detailed in Appendix D.5 and D.6. The following is a more detailed description of each cluster.

Cluster 1: Highly Positive Evaluators

Cluster 1, comprising 50% of the sample, predominantly exhibits positive evaluations across all indicators. Members of this cluster are very likely to agree with all the positively formulated statements, with agreement rates for statements ranging from 86% to 98%. Additionally, the disagreement with the negatively framed statement "I found the consultation difficult to understand" (76%) further underscores their positive evaluation. This cluster has a high average consultation rating of 8,09, reflecting strong overall satisfaction with the consultation. Demographically, this cluster consists of a balanced gender distribution (52% men and 48% women) and middle-aged individuals (48%). Most of the members have a medium (40%) education level and work full-time (60%). A majority live in urban areas (61%) and rarely face accessibility problems (71%).

Cluster 2: Skeptical Positive Evaluators

Cluster 2, which represents 30% of the sample, also tends to evaluate the PVE consultation positively but with a higher proportion of neutral responses compared to Cluster 1. Members of this cluster are more likely to disagree (88%) with the negatively framed statement that this consultation is difficult to understand, while agreeing with 3 of the positively formulated statements. The average rating for this cluster is quite high at 7,45, indicating a generally positive stance. This cluster features a balanced gender distribution (51% men and 49% women) and a higher representation of individuals between 35 and 65 years (51%). More than half of the individuals in this cluster have a high education level (50%) and work full time (62%).

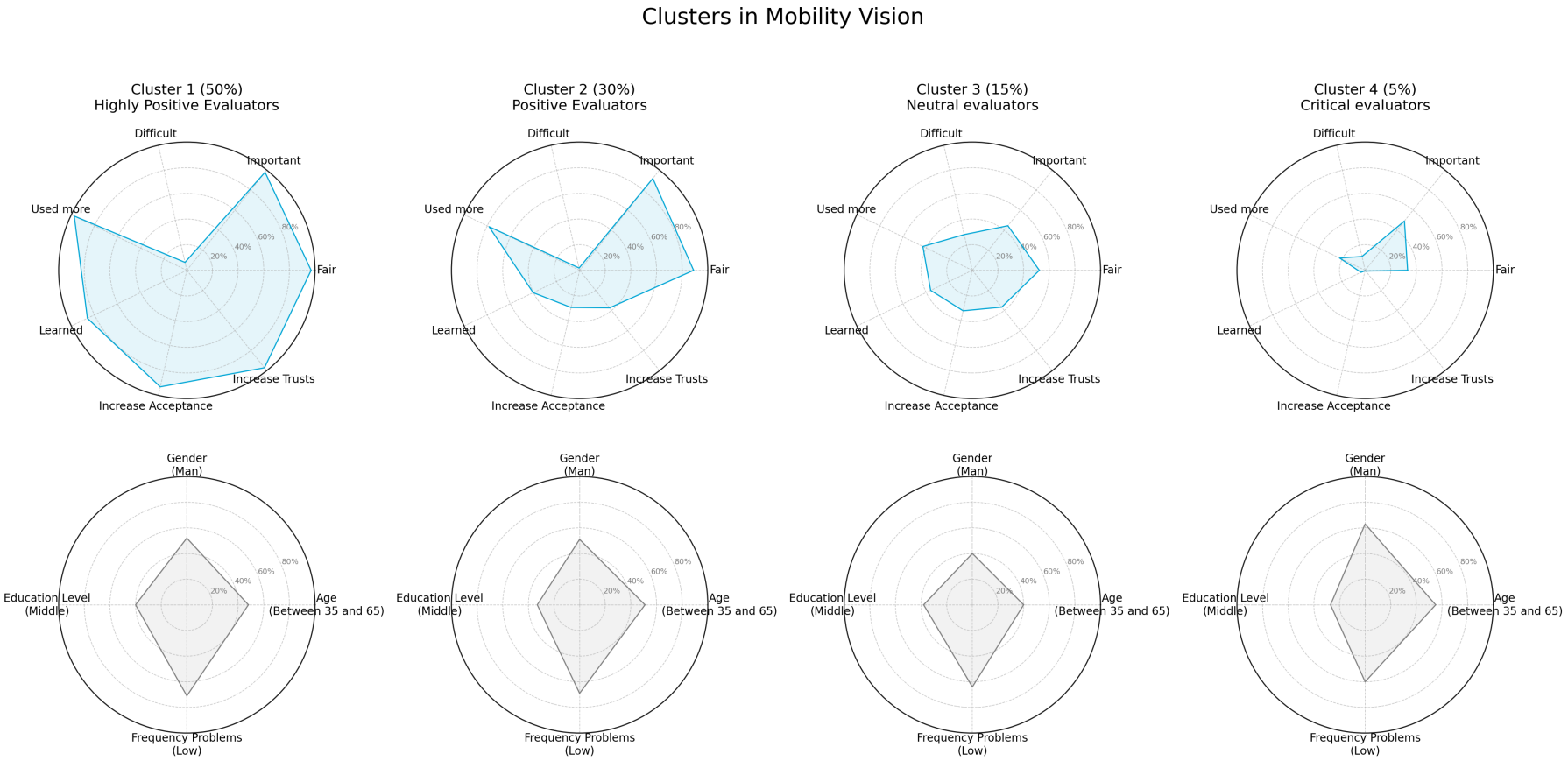


Figure 4.19: Clusters in Mobility Vision Dataset

Cluster 3: Neutral Evaluators

Cluster 3, which represents 15% of the sample, exhibits neutral responses to the face validity statements. Members of this cluster exhibit a higher tendency to remain impartial or have no strong opinion on four out of six statements that are phrased in a positive way. The members of this cluster exhibit the lowest likelihood of expressing disagreement with the statement ”This consultation is difficult to understand” compared to other clusters (43%). This cluster exhibits a comparatively lower average ratings of 6.46. From a demographic standpoint, individuals in this cluster have a significantly higher likelihood of being a young adult (41%) and being female (60%). This cluster also exhibits the largest likelihood of individuals with low levels of education in comparison to other clusters, which is consistent with their assessment of the difficulties associated with this consultation.

Cluster 4: Critical Evaluators

The smallest cluster, encompassing 5% of the sample, presents the most critical evaluations. Members of this cluster are more likely to disagree with three out of six positively framed indicators and being neutral with two of them. However, there is a notable proportion (67%) of disagreement with the negatively framed statement ”This consultation is difficult to understand”, which indicates some positive sentiment. The average rating for this cluster is the lowest at 5,93, reflecting more critical perspectives. Members of this cluster are more likely to be men (63%) and have a high education background (54%).

4.4.4 Comparison of Latent Class Cluster Analysis Across Consultations

The individual LCCA performed in each PVE consultation produced consistent outcomes, with 4 clusters identified in each consultation, as depicted in Figure 4.20. While the demographic characteristics of respondents in the corresponding clusters across different consultations (e.g., Cluster 1 in all datasets) have distinct features, their evaluations of face validity in each cluster are fairly comparable.

The largest group in each consultation typically gives the most positive assessment of the PVE consultation. Demographically, they primarily consist of middle-aged individuals and young adults who possess a medium level of education. They have a moderate probability of being train users; however, although they live closer to the project, they are less likely to use the line route in discussion.

The second largest clusters in all three PVE consultations likewise tend to have a neutral-positive evaluation of the consultation. Demographically, these individuals are predominantly middle-aged to seniors, with a medium to high educational background. The members of these clusters are likely to be non-workers. They are more likely to rarely use the train and the line in discussion.

The third largest clusters exhibit some variations between the three PVE consultations. Cluster 3 in the Lelylijn consultations is characterized by critical evaluators who are potential users of the line. In contrast, Cluster 3 in the Oude Lijn and Mobility Vision consultations includes neutral evaluators. In the Lelylijn consultation, members of Cluster 3 are more likely to be middle-aged people with medium education levels, not working, and residing at a distance from the project. In the Oude Lijn consultation, Cluster 3 members are predominantly middle-aged, full-time workers, with high education levels and residing close to the project. In the Mobility Vision consultation, Cluster 3 consists of young adults and middle-aged individuals, mostly women with low education levels.

Finally, the last and smallest clusters across the consultations tend to be the most critical in their evaluations. In the Lelylijn and Oude Lijn consultations, these clusters are critical evaluators who are non-users. In the Mobility Vision consultation, Cluster 4 also comprises critical evaluators. Demographically, the members of these clusters vary: in the Lelylijn consultation, Cluster 4 includes middle-aged individuals with high education levels, full-time workers, and infrequent train users. In the Oude Lijn consultation, Cluster 4 members are middle-aged and young adults, with low education levels and infrequent train users. In the Mobility Vision consultation, Cluster 4 consists of middle-aged and young adults, mostly men, with high education levels and infrequent train users. Both in the Oude Lijn and Lelylijn consultations, the members are more likely to live far from the project.

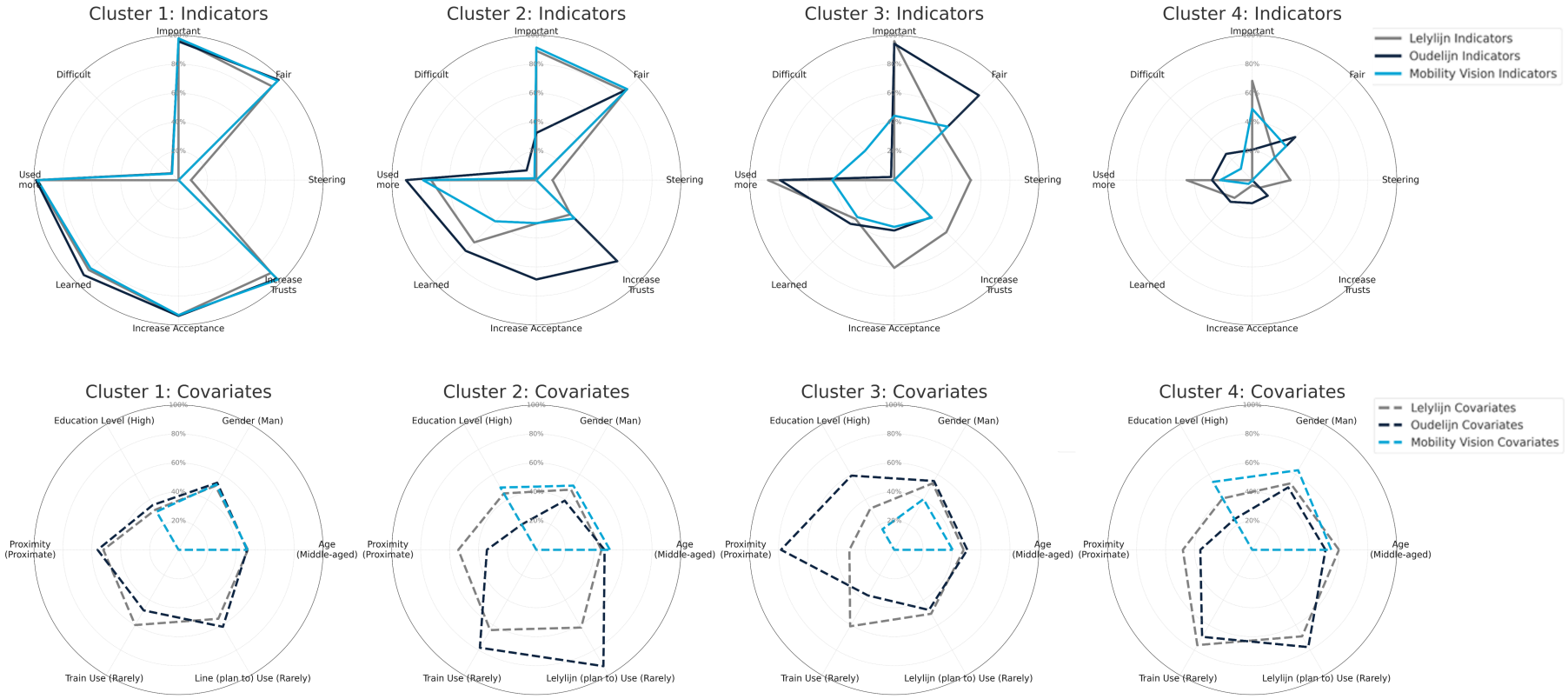


Figure 4.20: Clusters Comparisons in each Dataset

The comparisons of LCCA results between PVE consultations indicate that there are consistent segments that have emerged across all three consultations. The two most positive groups consist of middle-aged individuals and young adults with a moderate to high level of education. The cluster with the highest positive evaluation tends to consist of individuals who live in close proximity to the project and occasionally use the train. In contrast, the neutral-positive group, seldom make use of the train and plan to use the line. The two least positive clusters often comprise individuals ranging from young adults to middle-aged persons. Individuals belonging to these groups are more prone to possessing a lower level of educational level compared to the first two positive groups.

4.5 Multinomial Logistic Regression Results

To investigate the influence of respondent characteristics on the evaluation of each face validity statement individually, it was initially planned to conduct Ordinal Logistic Regression. This method was chosen due to the ordinal nature of the dependent variables, represented by Likert scale categories in the evaluation of each face validity statement. However, upon testing the parallel regression assumption, as outlined in Appendix E.1, it became evident that this assumption was not satisfied. Consequently, multinomial logistic regression was opted for instead.

The coding utilized in SPSS for this study is elaborated upon in Appendix F. Multinomial logistic regression was performed using weighted data for all consultations. However, the Pearson and Deviance statistics for each statement were found to be significant, indicating poor goodness-of-fit. Conversely, the non-weighted models, as provided in Appendix G, indicated at least one of the Pearson and Deviance statistics to be non-significant. Despite this, the weighted models were proceeded with for their better Pseudo R-square and correct percentage compared to the non-weighted models. Additionally, the Latent Class Cluster Analysis in the next section also showed better results for weighted data. Therefore, for the purpose of data consistency as well, the weighted data was used in all the analysis.

The following sections will present comprehensive explanations of the results of multinomial logistic regression for each face validity statement in the Lelylijn consultation. Subsequently, a comparison of the multinomial logistics regression outcomes from all three PVE consultations will be presented. In order to maintain brevity in this report, the comprehensive analysis of Oude Lijn and Mobility Vision are not discussed in this chapter, and can be accessed in Appendix G.2 and Appendix G.3 instead.

4.5.1 Lelylijn Project

The consultation steered my choices in a certain direction

The Multinomial Logistic Regression analysis was conducted for each face validity statement in Lelylijn consultation. The analysis was conducted with three outcome groups: (Totally) Agree, Neutral, and (Totally) Disagree. The regression was performed twice, using ”Neutral” and ”(Totally) Agree” as the reference outcomes, resulting in three comparisons: (Totally) Agree vs. Neutral, (Totally) Disagree vs. Neutral, and (Totally) Disagree vs. (Totally) Agree.

The ”B” column represents the estimated coefficients for the predictor variables. When the predictors are categorical, the coefficients indicate the change in the log-odds of the outcome categories relative to the reference category of the predictor variable. The ”Sig.” column in the results table indicates the significance level (p-value) for the test of the null hypothesis that the corresponding coefficient is equal to zero. If the p-value is less than 0,05, the null hypothesis is rejected, suggesting that the predictor variable significantly affects the outcome variable. Conversely, if the p-value is greater than or equal to 0,05, there is not enough evidence to suggest a significant effect of the predictor variable on the outcome variable.

The multinomial logistic regression conducted on the statement ”*The consultation steered my choices in a certain direction*” in the Lelylijn dataset was presented in Table 4.14. This statement is negatively formulated, hence the more probability of choosing (totally) disagree indicates positive evaluation, and vice versa. The analysis revealed 9 significant attributes when controlled for other independent variables.

Within this statement, it can be observed that:

1. **Age:**
   Middle-aged individuals (35-64 years) are less likely to (Totally) Disagree than (Totally) Agree (coefficient = -0,381, p = 0,016) compared to those 65 years or older. This indicates that middle-aged adults are more likely to think that this consultation influenced their choices, compared to seniors.
2. **Gender:**
   Men are more likely to (Totally) Agree than be Neutral (coefficient = 0,539, p < 0,001). They are also less likely to (Totally) Disagree than be Neutral (coefficient = -0,209, p = 0,004), and less likely to (Totally) Disagree than (Totally) Agree (coefficient = -0,749, p < 0,001) compared to women. This suggests that men are more likely to feel that the consultation has steered their choices in a certain direction.
3. **Education Level:**
   People with low education levels are less likely to (Totally) Agree than be Neutral (coefficient = -0,403, p = 0,002) and less likely to (Totally) Disagree than be Neutral (coefficient = -0,572, p < 0,001) compared to those with high education levels. Additionally, individuals with medium education levels are less likely to (Totally) Agree than be Neutral (coefficient = -0,339, p = 0,003) and more likely to (Totally) Disagree than (Totally) Agree (coefficient = 0,239, p = 0,03) compared to those with high education levels. This indicates that people with lower education levels are more neutral with whether this consultation steered their choices. On the other side, people with medium education level are less likely to feel that they are influenced by the consultation compared to those with higher education levels.
4. **Financial Status:**
   People who do not have enough money every month are more likely to (Totally) Agree than be Neutral (coefficient = 0,982, p < 0,001) and less likely to (Totally) Disagree than (Totally) Agree (coefficient = -0,698, p < 0,001) compared to those with more than enough money every month. People who have enough money are less likely to (Totally) Disagree than be Neutral (coefficient = -0,26, p = 0,001) and less likely to (Totally) Disagree than (Totally) Agree (coefficient = -0,302, p = 0,007) compared to those with more than enough money every month. This suggests that individuals who do not have enough money or just enough money every month are more likely to think that this consultation steered their choices into certain direction.
5. **Occupation:**
   People who are not working are less likely to (Totally) Disagree than be Neutral (coefficient = -0,36, p < 0,001) and less likely to (Totally) Disagree than (Totally) Agree (coefficient = -0,453, p < 0,001) compared to those working full-time. This indicates that unemployed individuals are more likely to think that their choices are being steered into certain direction.
6. **Residency:**
   People who live in someone's property are more likely to (Totally) Disagree than (Totally) Agree with the statement (coefficient = 0,381, p = 0,022). This indicates that people who live in someone's property are less likely to think that this consultation steered their choices, compared to people who live in an owner-occupied house.

Table 4.14: Results of multinomial logistic regression for statement: The consultation steered my choices in a certain direction

Attributes	Agree to Neutral		Disagree to Neutral		Disagree to Agree	
	B	Sig.	B	Sig.	B	Sig.
Age						
18-34 years	0,273	0,090	0,005	0,969	-0,269	0,094
35-64 years	0,284	0,075	-0,097	0,409	-0,381	0,016
65 years or older	REFERENCE					
Gender						
Man	0,539	<,001	-0,209	0,004	-0,749	<,001
Woman	REFERENCE					
Education						
Low	-0,403	0,002	-0,572	<,001	-0,169	0,193
Medium	-0,339	0,003	-0,099	0,234	0,239	0,030
High	REFERENCE					
Financial Status						
Have not enough money	0,982	<,001	0,284	0,096	-0,698	<,001
Have enough money	0,042	0,719	-0,26	0,001	-0,302	0,007
Have more than enough money	REFERENCE					
Occupation						
Not working	0,094	0,475	-0,360	<,001	-0,453	<,001
Part-time	-0,013	0,924	-0,041	0,671	-0,028	0,832
Full-time	REFERENCE					
Residency						
Live in a rental property	-0,069	0,584	-0,173	0,064	-0,104	0,400
Live in someone's property	-0,182	0,289	0,199	0,113	0,381	0,022
Live in an owner-occupied home	REFERENCE					
Proximity to the project						
Live not close to the project	0,014	0,883	0,154	0,027	0,140	0,128
Live close to the project	REFERENCE					
Frequency to use train						
Rarely	0,047	0,728	-0,468	<,001	-0,516	<,001
Occasionally	-0,352	0,019	-0,097	0,357	0,256	0,073
(Almost) daily	REFERENCE					
Frequency to use Lelylijn						
Rarely	-0,366	0,006	0,319	0,002	0,685	<,001
Occasionally	-0,573	<,001	0,106	0,313	0,679	<,001
(Almost) daily	REFERENCE					
Likelihood Ratio Test	<,001					
Pearson	<,001					
Deviance	<,001					
Pseudo R-Square: Cox and Snell	0,07					
Pseudo R-Square: Nagelkerke	0,08					
Pseudo R-Square: McFadden	0,036					
Correct Percentage	53,40%					

7. **Proximity to the Project:**
   People not close to the project are more likely to (Totally) Disagree than be Neutral (coefficient = 0,154, p = 0,027) compared to those living close to the project, suggesting that people who live far from the project are less likely to think that this consultation steered their choices.
8. **Frequency of Train Use:**
   People who rarely use trains are less likely to (Totally) Disagree than be Neutral (coefficient = -0,468, p < 0,001) and less likely to (Totally) Disagree than (Totally) Agree (coefficient = -0,516, p < 0,001) compared to those who use trains almost daily. This indicates that people who rarely use train are more like to think that this consultation steered their choices. Conversely, people who use train occasionally are less likely to (Totally) Agree with this statement compared to people who travel with train almost daily (coefficient = -0,352, p = 0,019). This indicates that occasional train users are less likely to think that this consultation is steering their choices.
9. **Frequency of Lelylijn Use:**
   People who rarely use Lelylijn are less likely to (Totally) Agree than be Neutral (coefficient = -0,366, p = 0,006), more likely to (Totally) Disagree than be Neutral (coefficient = 0,319, p = 0,002), and

more likely to (Totally) Disagree than (Totally) Agree (coefficient = 0,685, p < 0,001) compared to those who use it almost daily. Similarly, People who occasionally use Lelylijn are less likely to (Totally) Agree than being neutral (coefficient= -0,573, p < 0,001) and more likely to (Totally) Disagree than (Totally) Agree (coefficient = 0,679, p <0,001) with the statement, indicating they are less likely to think that this consultation steered their choices into certain direction, compared to individial who use Lelylijn (almost) daily.

The Likelihood Ratio Test was significant (p < 0,001), indicating the model's superiority over the intercept-only model in predicting whether the consultation steered respondents' choices. The Pseudo R-Square values (Cox and Snell, Nagelkerke, McFadden) suggest that the model explains a small proportion of the variance in the outcome variable. The model correctly classifies 53,4% of cases.

In conclusion, the analysis of whether respondents thought that the Lelylijn consultation steered their choices in a certain direction reveals several significant insights. Middle-aged people are less likely to (Totally) disagree with this statement, indicating that they tend to think the consultation indeed steered their choices. Men are more likely to (Totally) agree with this statement, suggesting that they feel their choices were influenced by the consultation.

People with low and medium education levels are less likely to (Totally) agree with this statement, indicating that they are less likely to think their choices were steered by the consultation compared to those with high education levels. Conversely, people who do not have enough money and those who have just enough money every month are less likely to (Totally) disagree with this statement, indicating that they are more likely to think their choices were steered by the consultation compared to people who have more than enough money every month.

Unemployed people are less likely to (Totally) disagree with this statement, suggesting that they are more likely to think their choices were steered by the consultation compared to those working full-time. In contrast, people who live in someone else's property are more likely to (Totally) disagree with this statement, indicating that they tend not to think the consultation steered their choices.

People who rarely use trains are more likely to think that the consultation steered their choices. Conversely, people who rarely and ocassionally travel through Lelylijn are more likely to think that the consultation did not steer their choices in a certain direction.

Overall, these findings suggest that demographic factors such as age, gender, education level, financial status, employment status, residency, and train usage frequency significantly influence whether respondents feel that the Lelylijn consultation steered their choices in a certain direction.

I trust this is a fair investigation

A second multinomial logistic regression was conducted with the second face validity statement, "*I trust this is a fair investigation*" serving as the dependent variable. The outcomes are detailed in Table 4.15, revealing 6 attributes that load significantly across one or more comparisons when they are controlled for other independent variables.

Within this statement, it can be observed that:

1. **Age:**
   Younger adults (18-34 years) are more likely to (Totally) Agree than be Neutral (coefficient = 0,902, p < 0,001), more likely to (Totally) Disagree than be Neutral (coefficient = 1,415, p < 0,001), and more likely to (Totally) Disagree than (Totally) Agree (coefficient = 0,512, p = 0,033) compared to those 65 years or older. This suggests that younger adults have polarized opinions about the fairness of the investigation.
2. **Education Level:**
   People with low education levels are more likely to (Totally) Agree than be Neutral (coefficient = 0,763, p < 0,001) and more likely to (Totally) Disagree than be Neutral (coefficient = 0,842, p < 0,001) compared to those with high education levels. People with medium education levels are more likely to (Totally) Agree than be Neutral (coefficient = 0,287, p = 0,007) compared to those with high education levels. This indicates that people with low and medium education levels have stronger opinions about the fairness of the investigation, either positively or negatively.

3. **Financial Status:**
   People who do not have enough money are less likely to (Totally) Disagree than be Neutral (coefficient = -0,881, p = 0,049) and less likely to (Totally) Disagree than (Totally) Agree (coefficient = -0,883, p = 0,024) compared to those with more than enough money. People who have enough money are less likely to (Totally) Agree than be neutral with this statement (coefficient = -0,483, p < 0,001). However, they are also less likely to (Totally) Disagree than be Neutral (coefficient = -0,668, p < 0,001) compared to those with more than enough money. This suggests that people with less financial resources are more likely to perceive the investigation as fair.

Table 4.15: Results of multinomial logistic regression for statement: I trust this is a fair investigation

Attributes	Agree to Neutral		Disagree to Neutral		Disagree to Agree	
	B	Sig.	B	Sig.	B	Sig.
Age						
18-34 years	0,902	<,001	1,415	<,001	0,512	0,033
35-64 years	-0,073	0,623	0,179	0,507	0,251	0,298
65 years or older	REFERENCE					
Gender						
Man	-0,09	0,354	0,152	0,359	0,242	0,093
Woman	REFERENCE					
Education						
Low	0,763	<,001	0,842	<,001	0,079	0,686
Medium	0,287	0,007	0,272	0,139	-0,014	0,930
High	REFERENCE					
Financial Status						
Have not enough money	0,002	0,993	-0,881	0,049	-0,883	0,024
Have enough money	-0,483	<,001	-0,668	<,001	-0,185	0,213
Have more than enough money	REFERENCE					
Occupation						
Not working	-0,060	0,656	-0,44	0,051	-0,380	0,049
Part-time	-0,126	0,315	-0,307	0,147	-0,181	0,326
Full-time	REFERENCE					
Residency						
Live in a rental property	-0,238	0,056	-0,036	0,859	0,202	0,246
Live in someone's property	-0,197	0,302	-0,282	0,345	-0,084	0,731
Live in an owner-occupied home	REFERENCE					
Proximity to the project						
Live not close to the project	-0,363	<,001	-0,298	0,056	0,064	0,635
Live close to the project	REFERENCE					
Frequency to use train						
Rarely	-0,544	<,001	0,467	0,054	1,011	<,001
Occasionally	0,097	0,556	0,573	0,033	0,477	0,034
(Almost) daily	REFERENCE					
Frequency to use Lelylijn						
Rarely	-0,068	0,644	-0,145	0,541	-0,076	0,700
Occasionally	-0,021	0,891	0,023	0,924	0,045	0,826
(Almost) daily	REFERENCE					
Likelihood Ratio Test	<,001					
Pearson	<,001					
Deviance	<,001					
Pseudo R-Square: Cox and Snell	0,061					
Pseudo R-Square: Nagelkerke	0,086					
Pseudo R-Square: McFadden	0,051					
Correct Percentage	80,30%					

4. **Occupation:**
   People who are not working are less likely to (Totally) Disagree than (Totally) Agree (coefficient = -0,380, p = 0,049) compared to those working full-time. This indicates that unemployed individuals are more likely to think that this consultation is fair.
5. **Proximity to the Project:**
   People not close to the project are less likely to (Totally) Agree than be Neutral (coefficient = -0,363, p < 0,001) compared to those living close to the project, suggesting that distance from the

project decreases positive perceptions of the investigation's fairness.

6. **Frequency of Train Use:**
   People who rarely use trains are less likely to (Totally) Agree than be Neutral (coefficient = -0,544, p < 0,001) and more likely to (Totally) Disagree than (Totally) Agree (coefficient = 1,011, p < 0,001) compared to those who use trains almost daily. Similarly, people who occasionally use trains are more likely to (Totally) Disagree than be Neutral (coefficient = 0,573, p = 0,033) and more likely to (Totally) Disagree than (Totally) Agree (coefficient = 0,477, p = 0,034) compared to those who use trains almost daily. This suggests that infrequent and occasional train users have a more negative perception of the investigation's fairness.

The Likelihood Ratio Test was significant (p < 0,001), indicating the model's superiority over the intercept-only model in predicting whether the consultation steered respondents' choices. The Pseudo R-Square values (Cox and Snell, Nagelkerke, McFadden) suggest that the model explains a small proportion of the variance in the outcome variable. The model correctly classifies 80,30% of cases.

In conclusion, the analysis of whether respondents trust that the Lelylijn consultation is a fair investigation reveals several significant insights. Younger adults have polarized opinions about the fairness of the investigation, indicating both strong agreement and disagreement compared to seniors. People with low education levels also have stronger opinions, either positively or negatively, compared to those with high education levels. While people with medium education level are more likely to think that this consultation is fair.

Financial status significantly impacts perceptions of fairness. Individuals with less financial resources are more likely to perceive the investigation as fair. Similar to unemployed people who are less likely to think that this consultation is not fair. People far from the project are less likely to see the investigation as fair compared to those living close to it.

Train usage frequency also plays a crucial role. Infrequent train users are less likely to perceive the investigation as fair. Similarly, occasional train users also have a more negative perception compared to those who use trains almost daily.

Overall, demographic factors such as age, education level, financial status, proximity to the project, and frequency of train use significantly influence whether respondents trust that the Lelylijn consultation is a fair investigation.

I thought it was an important topic to give my opinion on

The multinomial logistic regression conducted on the Lelylijn dataset for the statement "*I thought it was an important topic to give my opinion on*" revealing 7 attributes that load significantly across various comparisons when adjusted for other independent variables. The outcomes are summarized in Table 4.16.

Within this statement, it can be observed that:

1. **Age:**
   Younger adults (18-34 years) are more likely to (Totally) Agree than be Neutral (coefficient = 0,792, p < 0,001) and more likely to (Totally) Disagree than be Neutral (coefficient = 1,256, p = 0,008) compared to those 65 years or older. This suggests that younger adults find the topic important and are more polarized in their opinions compared to elderly.
2. **Gender:**
   Men are less likely to (Totally) Agree than be Neutral (coefficient = -0,816, p < 0,001) and more likely to (Totally) Disagree than (Totally) Agree (coefficient = 0,867, p = 0,002) compared to women. This indicates that men are less likely to think this topic is important compared to women.
3. **Education:**
   People with low education levels are more likely to (Totally) Agree than be Neutral (coefficient = 0,558, p = 0,002) compared to those with high education levels. People with medium education levels are more likely to (Totally) Agree than be Neutral (coefficient = 0,573, p < 0,001) and more likely to (Totally) Disagree than be Neutral (coefficient = 0,838, p = 0,013) compared to those with high education levels. This suggests that individuals with lower education levels find the

topic important, while people with medium education level are more polarized in their opinions compared to those with higher education levels.

4. **Financial Status:**
People who do not have enough money every month are less likely to (Totally) Agree than be Neutral (coefficient = -0,905, p = 0,002) compared to those with more than enough money every month. Similarly, people who have enough money every month are less likely to (Totally) agree with this statement compared to people who have more than enough money (coefficient = -0,378). This suggests that individuals with insufficient financial resources do not find the topic as important as much as people who have more than enough money every month.

Table 4.16: Results of multinomial logistic regression for statement: I thought it was an important topic to give my opinion on

Attributes	Agree to Neutral		Disagree to Neutral		Disagree to Agree	
	B	Sig.	B	Sig.	B	Sig.
Age						
18-34 years	0,792	<,001	1,256	0,008	0,464	0,270
35-64 years	0,248	0,251	0,237	0,609	-0,011	0,979
65 years or older	REFERENCE					
Gender						
Man	-0,816	<,001	0,051	0,873	0,867	0,002
Woman	REFERENCE					
Education						
Low	0,558	0,002	-0,136	0,763	-0,694	0,098
Medium	0,573	<,001	0,838	0,013	0,265	0,378
High	REFERENCE					
Financial Status						
Have not enough money	-0,905	0,002	-1,624	0,086	-0,719	0,429
Have enough money	-0,378	0,030	-0,273	0,410	0,105	0,717
Have more than enough money	REFERENCE					
Occupation						
Not working	-0,145	0,458	0,469	0,260	0,614	0,101
Part-time	-0,210	0,298	-0,359	0,455	-0,15	0,736
Full-time	REFERENCE					
Residency						
Live in a rental property	0,404	0,053	-1,388	0,012	-1,792	<,001
Live in someone's property	-0,453	0,071	-1,874	0,002	-1,420	0,009
Live in an owner-occupied home	REFERENCE					
Proximity to the project						
Live not close to the project	-0,309	0,028	0,808	0,014	1,117	<,001
Live close to the project	REFERENCE					
Frequency to use train						
Rarely	0,004	0,986	0,009	0,985	0,005	0,991
Occasionally	-0,113	0,626	0,558	0,239	0,671	0,112
(Almost) daily	REFERENCE					
Frequency to use Lelylijn						
Rarely	-1,149	<,001	0,303	0,647	1,452	0,016
Occasionally	-0,540	0,067	0,777	0,251	1,317	0,032
(Almost) daily	REFERENCE					
Likelihood Ratio Test	<,001					
Pearson	<,001					
Deviance	<,001					
Pseudo R-Square: Cox and Snell	0,045					
Pseudo R-Square: Nagelkerke	0,103					
Pseudo R-Square: McFadden	0,08					
Correct Percentage	93,00%					

5. **Residency:**
People who live in rental properties are less likely to (Totally) Disagree than be Neutral (coefficient = -1,388, p = 0,012) and less likely to (Totally) Disagree than (Totally) Agree (coefficient = -1,792, p < 0,001) compared to those who live in owner-occupied homes. Similarly, people who live in someone's property are less likely to (Totally) Disagree than be Neutral (coefficient = -1,874, p = 0,002) and less likely to (Totally) Disagree than (Totally) Agree (coefficient = -1,42, p = 0,009)

compared to those who live in owner-occupied homes. This suggests that individuals living in rental properties or someone else's property are more likely to find the topic important compared to those living in owner-occupied homes.

6. **Proximity to the Project:**
   People not close to the project are less likely to (Totally) Agree than be Neutral (coefficient = -0,309, p = 0,028), more likely to (Totally) Disagree than be Neutral (coefficient = 0,808, p = 0,014), and more likely to (Totally) Disagree than (Totally) Agree (coefficient = 1,117, p < 0,001) compared to those living close to the project. This suggests that people far from the project do not find the topic as important as people who live close to the project.
7. **Frequency to use Lelylijn:**
   People who rarely use Lelylijn are less likely to (Totally) Agree than be Neutral (coefficient = -1,149, p < 0,001) and more likely to (Totally) Disagree than (Totally) Agree (coefficient = 1,452, p = 0,016) compared to those who use it almost daily. Similarly, people who occasionally use Lelylijn are more likely to (Totally) Disagree than (Totally) Agree (coefficient = 1,317, p = 0,032) compared to those who use it almost daily. This suggests that infrequent users of Lelylijn do not find the topic as important compared to frequent users.

The Likelihood Ratio Test was significant (p < 0,001), indicating the model's superiority over the intercept-only model in predicting whether the consultation steered respondents' choices. The Pseudo R-Square values (Cox and Snell, Nagelkerke, McFadden) suggest that the model explains a small proportion of the variance in the outcome variable. The model correctly classifies 93,0% of cases.

In conclusion, the analysis whether the respondents thought that Lelylijn consultation was an important topic to give their opinion on reveals several significant insights. Younger adults find the topic important and are more polarized in their opinions compared to older adults. Men are less likely to think this topic is important compared to women. Individuals with lower education levels find the topic important, while people from medium education level are more polarized in their opinions compared to those with higher education levels.

People do not have enough money every month and people who have enough money every month are more likely to think that this consultation is not important, compared to people who have more than enough money every month. People who live in rental properties are more likely to think that this topic is important for them to give their opinion on, compared to people who live in owner-occupied house. Conversely, people far from the project do not find the topic as important. Finally, infrequent users of Lelylijn do not find the topic as important compared to frequent users.

Overall, demographic factors such as age, gender, education level, financial status, residency, proximity to the project, and frequency of Lelylijn use significantly influence whether respondents think the topic is important to give their opinion on.

By participating in this consultation, I learn about the choices the government has to make about the Lely Line

The multinomial logistic regression conducted on the Lelylijn dataset for the statement "*By participating in this consultation, I learn about the choices the government has to make about the Lely Line*" in Table 4.17, revealed 8 attributes that load significantly across various comparisons when controlled for other independent variables.

Within this statement, it can be observed that:

1. **Age:**
   Young adults (younger than 35 years old) are more likely to (Totally) Agree than being neutral with the statement (coefficient = 0,542, p <0,001), relative to seniors. Middle-aged individuals (35-64 years) are less likely to (Totally) Disagree than be Neutral (coefficient = -0,530, p = 0,011) and less likely to (Totally) Disagree than (Totally) Agree (coefficient = -0,451, p = 0,015) compared to those 65 years or older. This suggests that young adults and middle-aged individuals are more likely to think that they learned about the government's choices through this consultation compared to seniors.
2. **Gender:**
   Men are less likely to (Totally) Agree than be Neutral (coefficient = -0,431, p < 0,001) and less

likely to (Totally) Disagree than be Neutral (coefficient = -0,257, p = 0,049) compared to women. This suggests that men are more likely to have a neutral stance regarding learning about the government's choices from the consultation compared to women.

3. **Education:**
   People with a medium education level are more likely to (Totally) Agree than be Neutral (coefficient = 0,251, p = 0,008) compared to those with a high education level. This suggests that people with a medium education level are more likely to think that they learned about the government's choices through this consultation compared to those with a high education level.

Table 4.17: Results of multinomial logistic regression for statement: By participating in this consultation, I learn about the choices the government has to make about the Lely Line

Attributes	Agree to Neutral		Disagree to Neutral		Disagree to Agree	
	B	Sig.	B	Sig.	B	Sig.
Age						
18-34 years	0,542	<,001	0,314	0,137	-0,228	0,214
35-64 years	-0,079	0,549	-0,530	0,011	-0,451	0,015
65 years or older	REFERENCE					
Gender						
Man	-0,431	<,001	-0,257	0,049	0,174	0,127
Woman	REFERENCE					
Education						
Low	0,165	0,136	-0,107	0,546	-0,272	0,085
Medium	0,251	0,008	-0,004	0,980	-0,255	0,053
High	REFERENCE					
Financial Status						
Have not enough money	-0,169	0,367	-0,803	0,025	-0,634	0,052
Have enough money	-0,304	0,002	-0,160	0,273	0,144	0,254
Have more than enough money	REFERENCE					
Occupation						
Not working	0,209	0,066	-0,086	0,629	-0,295	0,061
Part-time	-0,014	0,896	-0,254	0,148	-0,240	0,125
Full-time	REFERENCE					
Residency						
Live in a rental property	0,076	0,487	-0,149	0,382	-0,225	0,129
Live in someone's property	-0,484	<,001	-0,443	0,051	0,041	0,836
Live in an owner-occupied home	REFERENCE					
Proximity to the project						
Live not close to the project	0,037	0,641	0,312	0,013	0,275	0,013
Live close to the project	REFERENCE					
Frequency to use train						
Rarely	-0,313	0,007	-0,199	0,270	0,114	0,470
Occasionally	0,311	0,019	0,476	0,015	0,165	0,309
(Almost) daily	REFERENCE					
Frequency to use Lelylijn						
Rarely	0,045	0,709	0,292	0,137	0,247	0,153
Occasionally	-0,063	0,610	0,512	0,009	0,576	<,001
(Almost) daily	REFERENCE					
Likelihood Ratio Test	<,001					
Pearson	<,001					
Deviance	<,001					
Pseudo R-Square: Cox and Snell	0,05					
Pseudo R-Square: Nagelkerke	0,063					
Pseudo R-Square: McFadden	0,032					
Correct Percentage	70,30%					

4. **Financial Status:**
   People who have enough money are less likely to (Totally) Agree than be Neutral (coefficient = -0,304, p = 0,002) compared to those with more than enough money. People who do not have enough money are less likely to (Totally) Disagree than be Neutral (coefficient = -0,803, p = 0,025) compared to those with more than enough money. This suggests people with enough money tend to be neutral rather than strongly agreeing, and those who do not have enough money also tend

to be neutral rather than strongly disagreeing about learning from the consultation.

5. **Residency:**
   Individual who live in someone's property are less likely to (Totally) Agree than being neutral with this statement (coefficient = -0,484, p < 0,001). This suggests that people who live in someone's property are less likely to think that they learned about the government's choices through this consultation.
6. **Proximity to the Project:**
   People not close to the project are more likely to (Totally) Disagree than be Neutral (coefficient = 0,312, p = 0,013) and more likely to (Totally) Disagree than (Totally) Agree (coefficient = 0,275, p = 0,013) compared to those living close to the project. This suggests that people far from the project are less likely to think that they learned about the government's choices through this consultation.
7. **Frequency to use train:**
   People who rarely use trains are less likely to (Totally) Agree than be Neutral (coefficient = -0,313, p = 0,007) compared to those who use trains almost daily. This indicates that individuals who rarely use trains are more likely to have a neutral opinion rather than strongly agreeing that they learned from the consultation. People who occasionally use trains are more likely to (Totally) Agree than be Neutral (coefficient = 0,311, p = 0,019) and more likely to (Totally) Disagree than be Neutral (coefficient = 0,476, p = 0,015) compared to those who use trains almost daily. This suggests that individuals who occasionally use trains are more likely to have a strong opinion, either positive or negative, about learning from the consultation.
8. **Frequency to use Lelylijn:**
   People who occasionally use Lelylijn are more likely to (Totally) Disagree than be Neutral (coefficient = 0,512, p = 0,009) and more likely to (Totally) Disagree than (Totally) Agree (coefficient = 0,576, p < 0,001) compared to those who use it almost daily. This suggests that occasional users of Lelylijn are less likely to think that they learned about the government's choices through this consultation, compared to daily users of Lelylijn.

The likelihood ratio test for this model yielded a significant result (p < 0,001), indicating its superiority over the intercept-only model in predicting whether the consultation influenced choices in a certain direction. Despite significant Pearson and Deviance statistics suggesting a lack of fit, this model correctly predicted 70,3% of choices in the statement and explained 6,3% of its variance.

In conclusion, the analysis on whether the respondents learned about the government's choices through the Lelylijn consultation reveals several significant insights. Young adults and middle-aged individuals are more likely to think that they learned about the government's choices through this consultation compared to seniors. Men are less likely to think that they learned about the government's choices through this consultation compared to women.

People with medium education levels are more likely to think that they learned about the government's choices through this consultation compared to those with high education levels. Individuals with sufficient financial resources are less likely to think that they learned about the government's choices through this consultation compared to those with more than enough money. People far from the project are less likely to think that they learned about the government's choices through this consultation. Occasional users of Lelylijn are less likely to think that they learned about the government's choices through this consultation, compared to the daily users.

In the Netherlands we should use this method more often to involve residents in government policy

The multinomial logistic regression conducted on the Lelylijn dataset for the statement "*In the Netherlands we should use this method more often to involve residents in government policy*" results in Table 4.18, revealed 7 attributes that load significantly across various comparisons when adjusted for other independent variables.

Within this statement, it can be observed that:

1. **Age:**
   Younger adults (18-34 years) are more likely to (Totally) Agree than be Neutral (coefficient = 0,436, p = 0,011), more likely to (Totally) Disagree than be Neutral (coefficient = 1,022, p = 0,002), and

more likely to (Totally) Disagree than (Totally) Agree (coefficient = 0,586, p = 0,043) compared to those 65 years or older. This indicates a polarized view among younger adults. While some younger adults believe strongly in using this method more often, others are strongly opposed to it, indicating a lack of consensus within this age group.

Table 4.18: Results of multinomial logistic regression for statement: In the Netherlands we should use this method more often to involve residents in government policy

Attributes	Agree to Neutral		Disagree to Neutral		Disagree to Agree	
	B	Sig.	B	Sig.	B	Sig.
Age						
18-34 years	0,436	0,011	1,022	0,002	0,586	0,043
35-64 years	0,023	0,887	0,041	0,902	0,018	0,952
65 years or older	REFERENCE					
Gender						
Man	-0,323	0,002	0,700	0,002	1,022	<,001
Woman	REFERENCE					
Education						
Low	0,758	<,001	0,462	0,094	-0,296	0,232
Medium	0,645	<,001	0,020	0,935	-0,624	0,005
High	REFERENCE					
Financial Status						
Have not enough money	0,439	0,085	0,816	0,038	0,377	0,238
Have enough money	-0,027	0,825	-0,786	<,001	-0,760	<,001
Have more than enough money	REFERENCE					
Occupation						
Not working	-0,306	0,031	-0,080	0,770	0,226	0,356
Part-time	-0,116	0,405	-0,496	0,128	-0,380	0,209
Full-time	REFERENCE					
Residency						
Live in a rental property	-0,907	<,001	-1,716	<,001	-0,810	0,003
Live in someone's property	-0,336	0,089	-1,105	0,002	-0,769	0,016
Live in an owner-occupied home	REFERENCE					
Proximity to the project						
Live not close to the project	0,149	0,139	-0,089	0,656	-0,238	0,182
Live close to the project	REFERENCE					
Frequency to use train						
Rarely	-0,351	0,017	-0,338	0,234	0,013	0,958
Occasionally	0,184	0,257	0,210	0,489	0,026	0,922
(Almost) daily	REFERENCE					
Frequency to use Lelylijn						
Rarely	-0,212	0,173	0,250	0,428	0,462	0,102
Occasionally	-0,232	0,145	0,134	0,675	0,366	0,202
(Almost) daily	REFERENCE					
Likelihood Ratio Test	<,001					
Pearson	<,001					
Deviance	<,001					
Pseudo R-Square: Cox and Snell	0,046					
Pseudo R-Square: Nagelkerke	0,073					
Pseudo R-Square: McFadden	0,048					
Correct Percentage	85,70%					

2. **Gender:**
   Men are less likely to (Totally) Agree than be Neutral (coefficient = -0,323, p = 0,002), more likely to (Totally) Disagree than be Neutral (coefficient = 0,7, p = 0,002), and more likely to (Totally) Disagree than (Totally) Agree (coefficient = 1,022, p < 0,001) compared to women. This suggests that men are less likely to think that this consultation should be used more often.
3. **Education Level:**
   People with medium education levels are more likely to (Totally) Agree than be Neutral (coefficient = 0,645, p < 0,001) and less likely to (Totally) Disagree than (Totally) Agree (coefficient = -0,624, p = 0,005) compared to those with high education levels. Similarly, people with low education level are more likely to (Totally) Agree with this statement than being neutral (coefficient = 0,758, p <

0,001), compared to people with high education background. This indicates that individuals with low and medium education levels are more supportive of using this method to involve residents, compared to people with high educaiton level.

4. **Financial Status:**
   People who have enough money are less likely to (Totally) Disagree than be Neutral (coefficient = -0,786, p < 0,001) and less likely to (Totally) Disagree than (Totally) Agree (coefficient = -0,760, p < 0,001) compared to those with more than enough money. This suggests that those with sufficient financial resources are more supportive of involving residents in government policy. Conversely, people who do not have enough money every month are more likely to think that this method should not be used more often (coefficient = 0,816, p = 0,038).
5. **Occupation:**
   People who are not working are less likely to (Totally) Agree than be Neutral (coefficient = -0,306, p = 0,031) compared to those working full-time. This indicates that unemployed individuals are less supportive of using this method to involve residents in government policy.
6. **Residency:**
   People who live in rental properties are less likely to (Totally) Agree than be Neutral (coefficient = -0,907, p < 0,001), less likely to (Totally) Disagree than be Neutral (coefficient = -1,716, p < 0,001), and less likely to (Totally) Disagree than (Totally) Agree (coefficient = -0,81, p = 0,003) compared to those living in owner-occupied homes. This suggests that renters may feel neutral about the involvement of residents in such consultation method for government policy. People who live in someone else's property are less likely to (Totally) Disagree than be Neutral (coefficient = -1,105, p = 0,002) and less likely to (Totally) Disagree than (Totally) Agree (coefficient = -0,769, p = 0,016) compared to those living in owner-occupied homes. This indicates a tendency towards a more positive stance on the issue.
7. **Frequency to Use Train:**
   People who rarely use trains are less likely to (Totally) Agree than be Neutral (coefficient = -0,351, p = 0,017) compared to those who use trains almost daily. This suggests that infrequent train users are less supportive of using this method to involve residents in government policy.

The likelihood ratio test for this model yielded a significant result (p < 0,001), indicating its superiority over the intercept-only model in predicting whether the consultation influenced choices in a certain direction. Despite significant Pearson and Deviance statistics suggesting a lack of fit, this model correctly predicted 85,7% of choices in the statement and explained 7,3% of its variance.

In conclusion, the analysis of the statement "In the Netherlands we should use this method more often to involve residents in government policy" reveals several significant insights. Younger adults are more likely to have strong opinions, either agreeing or disagreeing, compared to seniors. Men are less likely to think that this method should be used more, compared to women.

People with low and medium education levels are more likely to agree that this method should be used more often, compared to those with higher education. Individuals with sufficient financial resources are more likely to support the method, although people who do not have enough money every month are less likely to think so. Unemployed individuals are less likely to agree that this method should be used more often to involve residents in government policy. Residents are more likely to agree that this method should be used more, compared to homeowners. Lastly, people who rarely use trains are less likely to support the method.

If the government involves residents in choices on a large scale through this consultation, then the final decisions about the Lelylijn will be more acceptable to me

The multinomial logistic regression conducted on the Lelylijn dataset for the statement "*If the government involves residents in choices on a large scale through this consultation, then the final decisions about the Lelylijn will be more acceptable to me*" results in Table 4.19 revealed 8 attributes that load significantly across various comparisons when controlled for other independent variables.

Within this statement, it can be observed that:

1. **Age:**
   Younger adults (18-34 years) are more likely to (Totally) Agree than be Neutral (coefficient = 0,755,

p < 0,001), more likely to (Totally) Disagree than be Neutral (coefficient = 1,253, p < 0,001), and more likely to (Totally) Disagree than (Totally) Agree (coefficient = 0,498, p = 0,004) compared to those 65 years or older. Middle-aged individual also show similar results, where they are more likely to (Totally) Agree than be Neutral (coefficient = 0,596, p < 0,001), more likely to (Totally) Disagree than be Neutral (coefficient = 1,052, p < 0,001), and more likely to (Totally) Disagree than (Totally) Agree (coefficient = 0,456, p = 0,008) compared to those 65 years or older. The results indicate that younger adults and middle-aged people have polarized opinions on whether involving residents in decision-making makes the final decisions more acceptable, with strong views both in favor and against.

2. **Gender:**
   Men are more likely to (Totally) Agree than be Neutral (coefficient = 0,244, p = 0,002) and less likely to (Totally) Disagree than (Totally) Agree (coefficient = -0,239, p = 0,024) compared to women. This indicates that men are more likely increase their acceptance on the final decisions if residents are involved in the consultation.

Table 4.19: Results of multinomial logistic regression for statement: If the government involves residents in choices on a large scale through this consultation, then the final decisions about the Lely Line will be more acceptable to me

Attributes	Agree to Neutral		Disagree to Neutral		Disagree to Agree	
	B	Sig.	B	Sig.	B	Sig.
Age						
18-34 years	0,755	<,001	1,253	<,001	0,498	0,004
35-64 years	0,596	<,001	1,052	<,001	0,456	0,008
65 years or older	REFERENCE					
Gender						
Man	0,244	0,002	0,005	0,968	-0,239	0,024
Woman	REFERENCE					
Education						
Low	0,576	<,001	0,513	0,001	-0,063	0,639
Medium	0,151	0,093	-0,344	0,012	-0,494	<,001
High	REFERENCE					
Financial Status						
Have not enough money	0,042	0,815	-0,183	0,494	-0,225	0,332
Have enough money	0,073	0,416	-0,037	0,776	-0,110	0,341
Have more than enough money	REFERENCE					
Occupation						
Not working	0,171	0,132	0,485	0,003	0,315	0,022
Part-time	-0,184	0,078	-0,244	0,118	-0,060	0,667
Full-time	REFERENCE					
Residency						
Live in a rental property	0,136	0,198	0,551	<,001	0,414	0,001
Live in someone's property	-0,185	0,193	-0,164	0,438	0,021	0,910
Live in an owner-occupied home	REFERENCE					
Proximity to the project						
Live not close to the project	-0,326	<,001	-0,696	<,001	-0,371	<,001
Live close to the project	REFERENCE					
Frequency to use train						
Rarely	0,201	0,067	0,528	0,001	0,326	0,023
Occasionally	0,383	0,001	-0,133	0,475	-0,516	0,002
(Almost) daily	REFERENCE					
Frequency to use Lelylijn						
Rarely	-0,151	0,189	0,569	0,002	0,720	<,001
Occasionally	-0,308	0,009	0,399	0,032	0,707	<,001
(Almost) daily	REFERENCE					
Likelihood Ratio Test	<,001					
Pearson	<,001					
Deviance	<,001					
Pseudo R-Square: Cox and Snell	0,057					
Pseudo R-Square: Nagelkerke	0,069					
Pseudo R-Square: McFadden	0,034					
Correct Percentage	65,50%					

3. **Education Level:**
   People with low education levels are more likely to (Totally) Agree than be Neutral (coefficient = 0,576, p < 0,001) and more likely to (Totally) Disagree than be Neutral (coefficient = 0,513, p = 0,001) compared to those with high education levels. This suggests that individuals with lower education levels may have more intense reactions, either positively or negatively, to the consultation process compared to those with higher education levels. People with medium education levels are less likely to (Totally) Disagree than be Neutral (coefficient = -0,344, p = 0,012) and less likely to (Totally) Disagree than (Totally) Agree (coefficient = -0,494, p < 0,001) compared to those with high education levels. This indicates that people with medium education levels are more likely to increase their acceptance on the final decisions.
4. **Occupation:**
   People who are not working are more likely to (Totally) Disagree than be Neutral (coefficient = 0,485, p = 0,003) and more likely to (Totally) Disagree than (Totally) Agree (coefficient = 0,315, p = 0,022) compared to those working full-time. This suggests that unemployed individuals may not increase their acceptance on the final decisions about Lelylijn although more residents are involved, compared to people who work full time.
5. **Residency:**
   People who live in rental properties are more likely to (Totally) Disagree than be Neutral (coefficient = 0,551, p < 0,001) and more likely to (Totally) Disagree than (Totally) Agree (coefficient = 0,414, p = 0,001) compared to those living in owner-occupied homes. This indicates that renters may not accept to increase their acceptance on the final decisions better compared to homeowners.
6. **Proximity to the Project:**
   People not close to the project are less likely to (Totally) Agree than be Neutral (coefficient = -0,326, p < 0,001), less likely to (Totally) Disagree than be Neutral (coefficient = -0,696, p < 0,001), and less likely to (Totally) Disagree than (Totally) Agree (coefficient = -0,371, p < 0,001) compared to those living close to the project. This suggests that distance from the project correlates with a more neutral stance on the increase of acceptability on the final decisions made through the consultation.
7. **Frequency to Use Train:**
   People who rarely use trains are more likely to (Totally) Disagree than be Neutral (coefficient = 0,528, p = 0,001), and more likely to (Totally) Disagree than (Totally) Agree (coefficient = 0,326, p = 0,023) compared to those who use trains almost daily. This suggests that infrequent train users may not accept better the final decisions, compared to people who travel with trains daily. Conversely, people who occasionally use trains are more likely to (Totally) Agree than be Neutral (coefficient = 0,383, p = 0,001) and less likely to (Totally) Disagree than (Totally) Agree (coefficient = -0,516, p = 0,002) compared to those who use trains almost daily. This indicates that occasional train users are more likely to increase their acceptance on the final decisions.
8. **Frequency to Use Lelylijn:**
   People who rarely use Lelylijn are more likely to (Totally) Disagree than be Neutral (coefficient = 0,569, p = 0,002) and more likely to (Totally) Disagree than (Totally) Agree (coefficient = 0,72, p < 0,001) compared to those who use it almost daily. This suggests that infrequent Lelylijn users may not better accept the final decisions. Similarly, people who occasionally use Lelylijn are more likely to (Totally) Disagree than be Neutral (coefficient = 0,399, p = 0,032) and more likely to (Totally) Disagree than (Totally) Agree (coefficient = 0,707, p < 0,001) compared to those who use it almost daily. This indicates that infrequent and occasional Lelylijn users may also be less likely to increase their acceptance on the final decisions.

The likelihood ratio test for this model yielded a significant result (p < 0,001), indicating its superiority over the intercept-only model in predicting whether the consultation influenced choices in a certain direction. Despite significant Pearson and Deviance statistics suggesting a lack of fit, this model correctly predicted 65,5% of choices in the statement and explained 6,9% of its variance.

In conclusion, the analysis of the statement ”If the government involves residents in choices on a large scale through this consultation, then the final decisions about the Lely Line will be more acceptable to

me" reveals several significant insights. Younger and middle-aged adults are more likely to have strong and polarized opinions on whether involving residents in decision-making makes the final decisions more acceptable. Men are more likely to find the final decisions acceptable, while people with low and medium education levels also show more support for the process, compared to people from high education level.

Unemployed individuals and renters may have more reservations about the final decisions, indicating a potential lack of confidence in the process. People far from the project show neutral stance on the acceptability of the final decisions. Infrequent users of trains and Lelylijn are less likely to think that if the government involves residents in choices on a large scale through this consultation, it will increase their acceptance on the final decisions.

If the government allows residents to think about these types of choices more often, I will have more confidence in the government's decisions

The multinomial logistic regression conducted on the Lelylijn dataset for the statement "*If the government allows residents to think about these types of choices more often, I will have more confidence in the government's decisions*" results in Table 4.20, revealed 7 attributes that load significantly across various comparisons when controlled for other independent variables.

Within this statement, it can be observed that:

1. **Age:**
   Younger adults (18-34 years) are more likely to (Totally) Agree than be Neutral (coefficient = 0,583, $p < 0,001$) and more likely to (Totally) Disagree than be Neutral (coefficient = 0,79, $p < 0,001$) compared to those 65 years or older. Middle-aged individuals (35-64 years) are more likely to (Totally) Agree than be Neutral (coefficient = 0,334, $p = 0,005$) and more likely to (Totally) Disagree than be Neutral (coefficient = 0,499, $p = 0,018$) compared to those 65 years or older. This indicates that middle-aged individuals, like younger adults, tend to have stronger and more divided views on the matter of trust on the government, if they allow residents to think about these types of choices through PVE.
2. **Gender:**
   Men are more likely to (Totally) Agree than be Neutral (coefficient = 0,271, $p < 0,001$) and more likely to (Totally) Disagree than be Neutral (coefficient = 0,373, $p = 0,004$) compared to women. The findings indicate that men are more polarized in their opinions compared to women regarding the statement that experience that participating in PVE increases their trust in the government.
3. **Education Level:**
   People with low education levels are more likely to (Totally) Agree than be Neutral (coefficient = 0,302, $p = 0,003$), less likely to (Totally) Disagree than be Neutral (coefficient = -0,516, $p = 0,003$), and less likely to (Totally) Disagree than (Totally) Agree (coefficient = -0,818, $p < 0,001$) compared to those with high education levels. Similarly, people with medium education levels are more likely to (Totally) Agree than be Neutral (coefficient = 0,464, $p < 0,001$), less likely to (Totally) Disagree than be Neutral (coefficient = -0,463, $p = 0,002$), and less likely to (Totally) Disagree than (Totally) Agree (coefficient = -0,927, $p < 0,001$) compared to those with high education levels. This indicates that people with low and medium education levels experience that participating in PVE increases their trust in the government
4. **Financial Status:**
   People who do not have enough money are more likely to (Totally) Disagree than be Neutral (coefficient = 1,145, $p < 0,001$) and more likely to (Totally) Disagree than (Totally) Agree (coefficient = 0,874, $p < 0,001$) compared to those with more than enough money. These findings indicate that individuals with insufficient financial resources are less inclined to believe that engaging in PVE activities enhances their trust in the government. In contrast, those who possess sufficient monthly income are more inclined to perceive that engaging in PVE activities enhances their trust in the government.
5. **Residency:**
   People who live in someone's property are less likely to (Totally) Disagree than be Neutral (coefficient = -0,895, $p < 0,001$) and less likely to (Totally) Disagree than (Totally) Agree (coefficient = -0,741, $p = 0,004$) compared to those living in an owner-occupied home. This indicates that

people who live in someone's property are more likely to experience that participating in PVE increases their trust in the government.

Table 4.20: Results of multinomial logistic regression for statement: If the government allows residents to think about these types of choices more often, I will have more confidence in the government's decisions

Attributes	Agree to Neutral		Disagree to Neutral		Disagree to Agree	
	B	Sig.	B	Sig.	B	Sig.
Age						
18-34 years	0,583	<,001	0,790	<,001	0,207	0,304
35-64 years	0,334	0,005	0,499	0,018	0,165	0,404
65 years or older	REFERENCE					
Gender						
Man	0,271	<,001	0,373	0,004	0,102	0,395
Woman	REF					
Education						
Low	0,302	0,003	-0,516	0,003	-0,818	<,001
Medium	0,464	<,001	-0,463	0,002	-0,927	<,001
High	REFERENCE					
Financial Status						
Have not enough money	0,271	0,130	1,145	<,001	0,874	<,001
Have enough money	0,172	0,041	-0,010	0,941	-0,183	0,164
Have more than enough money	REFERENCE					
Occupation						
Not working	0,009	0,935	0,009	0,959	0,001	0,997
Part-time	0,033	0,745	0,211	0,194	0,178	0,234
Full-time	REFERENCE					
Residency						
Live in a rental property	-0,105	0,280	-0,069	0,669	0,036	0,811
Live in someone's property	-0,154	0,250	-0,895	<,001	-0,741	0,004
Live in an owner-occupied home	REFERENCE					
Proximity to the project						
Live not close to the project	-0,106	0,147	-0,181	0,142	-0,075	0,513
Live close to the project	REFERENCE					
Frequency to use train						
Rarely	-0,201	0,057	0,449	0,016	0,650	<,001
Occasionally	0,169	0,136	0,430	0,031	0,260	0,155
(Almost) daily	REFERENCE					
Frequency to use Lelylijn						
Rarely	-0,079	0,466	0,326	0,084	0,406	0,020
Occasionally	-0,061	0,583	0,006	0,977	0,067	0,714
(Almost) daily	REFERENCE					
Likelihood Ratio Test	<,001					
Pearson	<,001					
Deviance	<,001					
Pseudo R-Square: Cox and Snell	0,052					
Pseudo R-Square: Nagelkerke	0,064					
Pseudo R-Square: McFadden	0,032					
Correct Percentage	65,00%					

6. **Frequency to Use Train:**
   People who rarely use trains are more likely to (Totally) Disagree than be Neutral (coefficient = 0,449, p = 0,016) and more likely to (Totally) Disagree than (Totally) Agree (coefficient = 0,65, p < 0,001) compared to those who use trains almost daily. Similarly, people who occasionally use trains are more likely to (Totally) Disagree than be Neutral (coefficient = 0,43, p = 0,031) compared to those who use trains almost daily. This suggests that infrequent and occasional train users are less likely to experience that participating in PVE increases their trust in the government.
7. **Frequency to Use Lelylijn:**
   People who rarely use Lelylijn are more likely to (Totally) Disagree than (Totally) Agree (coefficient = 0,406, p = 0,02) compared to those who use it almost daily. This suggests that infrequent Lelylijn users are less likely to experience that participating in PVE increases their trust in the government.

The likelihood ratio test for this model yielded a significant result (p < 0,001), indicating its superiority

over the intercept-only model in predicting whether the consultation influenced choices in a certain direction. Despite significant Pearson and Deviance statistics suggesting a lack of fit, this model correctly predicted 65% of choices in the statement and explained 6,4% of its variance.

In conclusion, the analysis of the statement ”If the government allows residents to think about these types of choices more often, I will have more confidence in the government’s decisions” reveals several significant insights. Younger people, middle-aged adults, and men, have polarized view regarding their confidence in the government’s decisions if residents are involved more often.

People who do not have enough money, infrequent train users, and infrequent Lelylijn users are less likely to have confidence in the government’s decisions if residents are involved more often. Conversely, people with low and medium education levels, people who have enough money every month, and people who live in someone’s property are more likely to have confidence in the government’s decisions about Lelylij, if the government allows residents to think about these types of choice more often.

4.5.2 Comparison of Multinomial Logistic Regression Across Consultations

The individual Multinomial Logistic Regression for each PVE consultation reveals numerous significant effects of the characteristics of the respondents to their assessment of the face validity statements in each consultation. However, there are instances where this trend is consistent throughout consultations and instances where it is not. The comparisons of these relationships across three PVE consultations are resulted as follows:

1. **Age**
   Negative: none.
   Positive: Older individuals consistently rate the fairness of the consultation, the ease of consultation, the importance of the topic, and the increase of acceptance on the final decisions in a positive manner.
   Inconsistency: There is no consistent and significant correlation between someone’s age and their opinion of whether they learned from the consultation, whether the approach should be used more often, and the increase of trust in government’s conclusions.

2. **Gender**
   Negative: Men consistently exhibit lower levels of perceived fairness, higher levels of perceived difficulty, lower levels of perceived information acquisition, lower levels of support for the utilization of this method, and lower levels of trust in the government’s decisions.
   Positive: none.
   Inconsistency: There is no consistent influence of an individual’s gender and their assessment of the importance of a given topic or their inclination to endorse governmental policies.

3. **Education**
   Negative: There is a persistent pattern indicating that those with higher levels of education are less inclined to learn, increase their acceptance, and trust in the government.
   Positive: The higher education someone has, the less likely they think the consultation is difficult.
   Inconsistency: There is no consistent pattern whether someone’s education level affect their evaluation on the fairness and the importance of the PVE consultation.

4. **Occupation**
   Negative: Despite the consistency cannot be checked, people who work full-time are less likely to think that the consultation is honest.
   Positive: Although the consistency cannot be verified, individuals who are employed on a full-time basis are less inclined to perceive the consultation as difficult.
   Inconsistency: none.

5. **Residency**
   Negative: Despite consistency cannot be checked, people who live in a rental house are more

likely to think that this consultation is difficult to understand, compared to home-owners.
Positive: People who lives in someone's property are less likely to think this consultation is steering.
Inconsistency: There is no consistent correlation between the kind of house someone lives in and the importance or frequency of PVE consultations.

6. **Proximity**
   Negative: Despite the consistency cannot be checked, people who lives far from the project are more likely to think this consultation is difficult.
   Positive: There is a consistent influence of someone's proximity with a project with the more important they think the PVE consultation is.
   Inconsistency: It is uncertain whether an individual's proximity to a project influences their perception of whether they have gained knowledge through consultation or not.

7. **Frequency of Train Use**
   Negative: none.
   Positive: There is a consistent pattern that people who travel with train more frequently tend to think this this is an honest investigation and be confident about the government's decisions. Additionally, the more frequent someone use train, the less likely they think the consultation is difficult or steering. They also tend to think the topic is important.
   Inconsistency: It is uncertain whether train use frequency influences their perception of whether they have gained knowledge through consultation or not.

8. **Frequency of (planned) Line Use**
   Negative: Although the consistency cannot be checked, people who (plan to) use the line more frequently are more likely to think that this consultation is steering.
   Positive: There is a consistent pattern that people who (plan to) use the line more frequently are more likely to think that the topic is important and to increase their acceptance on the final decisions.
   Inconsistency: It is uncertain whether line (planned) use frequency influences their perception of whether they have gained knowledge through consultation or not.

9. **Live/ work near stations**
   Negative: Although the consistency cannot be verified, there is a noticeable trend among individuals residing or working in close proximity to the listed stations. They are more inclined to perceive the consultation as difficult and less likely to believe that they gain knowledge from it or that it should be used more.
   Positive: none.
   Inconsistency: none.

10. **Get in/out at stations**
    Negative: Although the consistency is impossible to confirm, individuals who have previously used the mentioned stations are more like to see the consultation as difficult and less likely to consider the topic as important to give their opinion on.
    Positive: Although the consistency cannot be verified, individuals who have previously used the mentioned stations are more inclined to perceive this consultation as genuine, derive knowledge from it, and advocate for its more frequent implementation.
    Inconsistency: none.

11. **Area of Living**
    Negative: none.
    Positive: Although the consistency cannot be verified, urban dwellers are more likely to think that the topic is important, that they learned from the consultation and that this should be done more often.

Inconsistency: It is uncertain whether there is consistent influence of whether someone live in urban or rural area, with their evaluation is difficult and on they increase their acceptance the final decisions.

12. **Financial Status**
    Negative: none.
    Positive: Despite no consistent evidence, people who do not have money every month are more likely to think this is an honest investigation.
    Inconsistency: none.

13. **Frequency of Facing Accessibility Problems**
    Negative: Despite no consistent evidence, people who frequently face accessibility problems are less likely to think that this topic is important, learned from the consultation, think that this should be done more often and to increase their acceptance on the final decisions.
    Positive: none.

4.6 Summary

Evaluation of Face Validity in PVE Consultations

This chapter presents a quantitative evaluation of face validity in three PVE consultations: Lelylijn, Oude Lijn, and Mobility Vision. The analysis focuses on understanding how different demographic groups assess the overall and face validity of these consultations, and how consistent these relationships are, answering **sub-question 1.2 and 1.3.**

- **Descriptive statistics** show generally positive evaluations of the overall consultations. However, there is room for improvement in aspects of face validity, particularly in perceptions of the consultations being steering, difficult, and their role in increasing citizens' acceptance of final decisions and trust in government.
- **Crosstabulation analysis** reveals significant and generally consistent associations between respondents' age, gender, education level, proximity, and frequency of train and line use with their evaluations of face validity.
- **Latent Class Cluster Analysis (LCCA)** identifies four distinct participant segments in each consultation. The clusters are generally comparable across consultations, with the most positive evaluators being the largest group in all cases.
- **Multinomial Logistic Regression** indicates that while men, young adults, and highly educated individuals tend to rate the overall consultation highly, they often give lower evaluations on certain aspects of PVE face validity.
- Some respondent characteristics, such as residency, living or working near stations, frequency of station use, area of living, financial status, and frequency of facing accessibility problems, were not found to be significant or consistently associated with evaluations of PVE face validity.

5 Underlying Reasons for Negative Evaluations on PVE Consultations

The series of quantitative analysis performed in Chapter 4 revealed associations between several characteristics of the respondents and their evaluation of face validity in three separate PVE consultations. Within all three consultations, there are consistently one or more negative clusters that assess the PVE consultation in an unfavorable manner. In order to gain a more thorough understanding of the reasons why respondents were critical of PVE consultation, a content analysis was conducted to examine the elements that influenced their judgment. More precisely, the analysis concentrated on finding the reasons behind the dislike of the Lelylijn consultation. This knowledge is essential for making improvements in future PVE consultations. This chapter will provide answers to research question 2: *"What are the factors contributing to the negative assessment of PVE by these specific segments?"*. This chapter begins by presenting the results of content analysis of the whole Lelylijn consultation dataset, per each face validity dimension. Subsequently, the analysis of the primary factors among the categories of respondents who provided the most comments during the Lelylijn consultation is presented.

5.1 Content Analysis on Overall Lelylijn Dataset

The results of the content analysis regarding motivations for dislikes are presented in Figure 5.1, with details in Appendix H. Although the analysis encompassed the entire actual (non-weighted) dataset of 4.691 respondents, filtering was applied to exclude empty comments, unrelated comments (such as those pertaining to Lelylijn project in general rather than the consultation), and positive feedback on the consultation. Ultimately, the analysis focused on 2.084 responses. These responses were categorized into five dimensions of face validity. Notably, the most comments were about the Feasibility with 724 people commented about it, followed by Completeness (540), Acceptance (507), Transparency (226), and Clarity (87).

The Feasibility dimension primarily encompasses comments on the ease of making choices within the consultation process. Notably, in the context of the Lelylijn project consultation, which includes both a Participatory Value Evaluation (PVE) and a Discrete Choice Experiment (DCE) segment, criticisms were frequently directed towards the DCE segment. Although most comments are on the general consultation, there are also comments that particularly focus on the PVE and DCE components individually. Significantly, the DCE part received numerous mentions of dislikes (347), whereas dissatisfaction of the PVE part was less prevalent but still notable (67). Furthermore, participants frequently highlighted the consultation's time-consuming and lengthy character (173). The consultation experience (55) also highlighted practical challenges such as the difficulty of answering and typing long responses on cellphones, as well as the inability to navigate back to previous pages.

The completeness dimension primarily concerns the relevant data that is accessible to individuals for making well-informed decisions. A significant number of individuals have expressed concerns with the lack of information regarding the subsequent actions following the consultation (272). Specifically, they are interested in knowing what the government intends to do with the results and have expressed a desire to be informed about the outcomes and follow-up of the consultation. In addition, they expressed a desire for more information regarding the Lelylijn project as a whole (200), including updates on its current status and a thorough assessment of its potential effects.

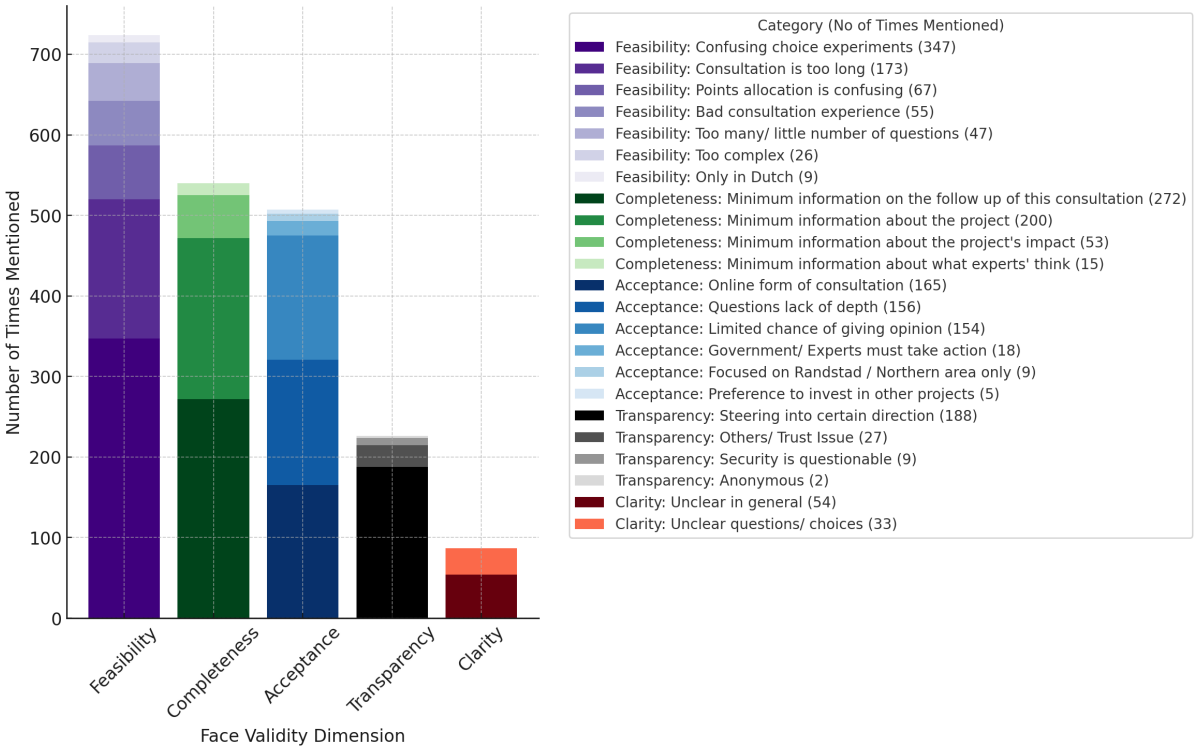


Figure 5.1: Disliked Aspects of the Lelylijn Consultation According to Respondents

The acceptance dimension corresponds to the effectiveness of this consultation as a means of providing advise to the government and the level of confidence respondents have in their choice. A total of 165 individuals expressed concerns about the online format of this consultation, highlighting the requirement for digital literacy and access to digital devices in order to participate and voice their opinions on this issue. A substantial proportion of individuals noted that the questions or scenarios presented in the consultation exhibit a deficiency in thoroughness (156). Furthermore, a significant number of individuals expressed limited opportunities to freely express their genuine viewpoints regarding the topics or options presented (154).

The transparency dimension refers to the extent to which the respondents perceived this consultation to be trustworthy and whether they felt that the research was guiding them towards a specific outcome. 188 individuals expressed that this consultation influenced their perspective, indicating that the construction of Lelylijn is inevitable. Several individuals expressed their distrust in the government, suggesting that this type of consultation may be employed to bolster predetermined decisions rather than genuinely seeking input from the population. Other comments express skepticism about the value of such consultations, as they believe the government always acts according to its own agenda regardless. Additionally, concerns were raised regarding the website's security and the possibility of individuals submitting multiple responses to this anonymous consultation. Finally, 87 respondents expressed their dissatisfaction with the consultation process due to its lack of clarity. The questions or the overall consultations were judged to be unclear.

5.2 Comments Given numerous times by "Negative" Evaluators Group

The quantitative analysis conducted in Chapter 4 yielded consistent negative evaluations of face validity of PVE from specific demographic groups. These demographic groups include young adults, males, individuals with a high level of education, people who live distant from the project, and individuals who infrequently use trains and have intentions to use them less frequently. In order to have a deeper understanding of the problems faced by these groups, the findings of the analysis of the content are further elaborated for each group of negative evaluators.

Figure 5.2 and Table 5.1 exhibit the comments that were most frequently stated by the groups of evaluators with negative opinions.

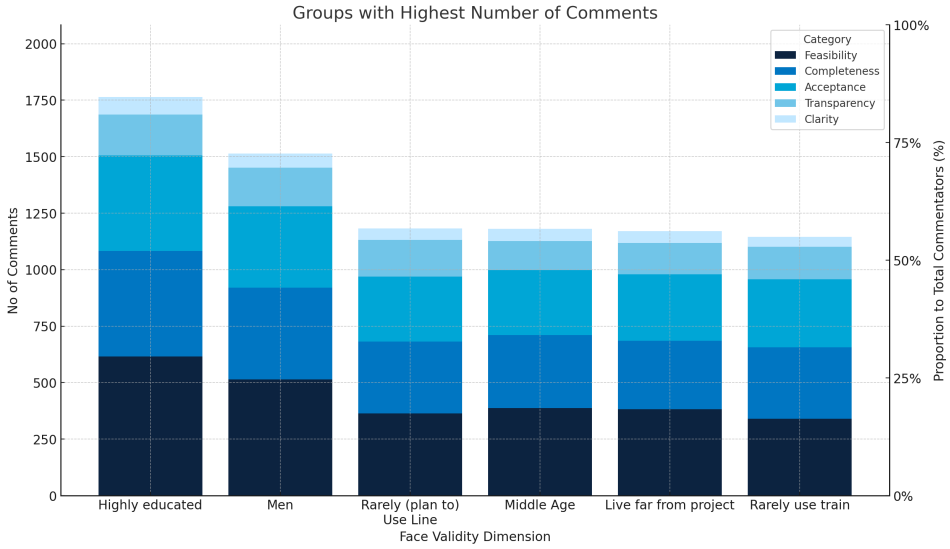


Figure 5.2: Groups with Highest Number of Comments per Face Validity Dimension

Individuals with a high level of education

The demographic that mentions critics most frequently is Highly educated Individuals. Out of the 2.084 comments submitted, 1.764 were provided by individuals with a high level of education. The statement that was most frequently made within this group pertained to the confusing nature of the choice experiments (296 mentions). This statement is contradictory as the quantitative analysis reveals that individuals with lower levels of education are more inclined to perceive the consultation as difficult. This could be attributed to the fact that the confusing choice experiments were the most frequently stated comment, independent of the respondents' characteristics. The second and third most frequently cited remarks among individuals with a high level of education pertain to the minimal amount of information they obtained during the consultation (231 mentions) and the details on the project (174 mentions). These are the possible factors that could account for the findings of the quantitative study, which indicate that those with greater levels of education are less inclined to believe that they have acquired knowledge about the government's choices, as well as to increase their acceptance and trust in the government.

Men

With 1.514 unfavorable comments given during the PVE consultation, men are the second most frequent critics. Similar to highly educated individuals, men provided the most comments on confusing choice experiments (255 mentions), insufficient information on the follow-up of consultation (197 mentions), and inadequate information about the project (157 mentions). This finding supports the quantitative analysis results that indicate that men are less inclined to perceive this consultation as being honest, learn knowledge about the government's decisions, and exhibit acceptance and confidence in the government's choices. Additionally, they are more inclined to see the consultation as difficult. In addition, the fourth most frequently stated comment, with 147 mentions, also aligns with the quantitative analysis findings indicating men are more likely to believe that this consultation is steering into a certain direction.

Table 5.1: Negative Comments in Lelylijn Consultation and their Frequency of Mentions

No	Comment	Face Validity Dimension	Frequency Mentioned					
			Highly educated	Men	Rarely (plan to) use the line	Live far from project	Rarely use train	Young adults
1	Confusing choice experiments	Feasibility	296	255	172	189	162	96
2	Minimum information on the follow up of consultation	Completeness	231	197	160	152	177	57
3	Steering into certain direction	Transparency	154	147	133	117	114	34
4	Minimum information about the project	Completeness	174	157	106	107	97	57
5	Consultation is too long	Feasibility	144	114	84	95	87	55
6	Online form of consultation	Acceptance	149	107	87	90	103	39
7	Limited chance of giving opinion	Acceptance	125	108	99	87	87	38
8	Questions lack of depth	Acceptance	120	120	88	96	94	29
9	Points allocation is confusing	Feasibility	60	46	32	33	27	33
10	Minimum information about the project's impact	Completeness	48	39	44	37	36	9
11	Unclear in general	Clarity	47	44	35	33	32	8
12	Bad consultation experience	Feasibility	49	40	30	26	28	17
13	Too many/ little number of questions	Feasibility	34	35	31	21	24	13
14	Unclear questions/ choices	Clarity	30	18	17	19	12	6
15	Others/ Trust Issue	Transparency	20	17	21	17	22	5
16	Too complex	Feasibility	23	20	10	13	8	4
17	Government/ Experts must take action	Acceptance	16	14	9	14	10	7
18	Minimum information about what experts' think	Completeness	14	12	8	6	6	4
19	Focused on Randstad / Northern area only	Acceptance	9	9	5	3	7	1
20	Security is questionable	Transparency	6	6	5	5	6	2
21	Only in Dutch	Feasibility	9	5	4	6	4	3
22	Preference to invest in other projects	Acceptance	4	2	1	4	-	-
23	Anonymous	Transparency	2	2	2	-	2	-
Grand Total			**1.764**	**1.514**	**1.183**	**1.170**	**1.145**	**517**

Individuals who rarely (plan to) Use Lelylijn

The third largest demographic that expressed negative feedback consists of individuals who infrequently (intend to) go via Lelylijn. The quantitative research revealed that this group has less of a tendency to consider this topic to be important for expressing their viewpoint and to accept the government's decisions. Although the majority of comments in this group fell into the same categories as the two groups outlined before, there were explicit comments made by individuals who are not potential users of Lelylijn. These comments may offer an explanation for the poor ratings. Quoting a resident, "I live in Eindhoven, so my opinion is not really of great importance in terms of participation in this decision-making". These individuals reside far from Lelylijn and hence have a lower likelihood of intending to utilize Lelylijn. Consequently, they believe that the subject lacks importance, thus making it unnecessary for them to express their opinion on.

One notable category in which this group surpasses the preceding two major groups is the matter of trust issues, with a total of 21 mentions. "I have the feeling that the decision about the Lely line has already been made. This without proper consultation with the people who are most affected by it," as stated by an individual. In addition, with regards to the minimum information on follow-up of this consultation, someone has commented that it is not completely evident how this consultation would be taken into account in the decision-making process. This could be attributed to the less accepted the choices made by the government by this group.

Individuals who Live far from Lelylijn

Individuals residing far from Lelylijn are among the most negative groups when it comes to the PVE consultations. The quantitative research revealed that this group has a lower likelihood of perceiving the topic as important. One potential reason is due to the virtual nature of this consultation which enables the participation of those who may not have a direct impact on the project. This is evident in the numerous references to the online method of consultation (87 mentions). An exemplary comment from this area was "Opinions of residents who will live close to the line are of the utmost importance. City dwellers cannot make good decisions about villages and countryside".

Another finding from the quantitative study revealed that this group is more inclined to see this consul-

tation as difficult. This speaks to the often expressed remarks that are made in a different way, such as ”It is too long and complex. It is probably difficult for many people to complete/assess properly”. Similar comments also found and fall into the category of consultation is too long (with 95 mentions) and confusing (with 189 mentions).

Individuals who rarely Use Train

Infrequent train users exhibit a negative perspective in multiple statements in PVE consultations. They are less likely to perceive the consultation as honest, less inclined to perceive the topic to be important, and less inclined to accept and have confidence in the government’s decisions. This phenomenon is evident in the substantial amount of comments relating to the insufficient details regarding the follow-up of the consultation (177) and the fact that the consultation is being guided towards a specific direction (114). One comment that accurately summarizes the above judgment is, ”The questions are very guiding. The entire setting of the consultation already assumes that the line will be built anyway. Can’t really be called objective. It is also poorly publicized that this consultation is taking place, probably with the part that mainly proponents will complete it.”

Another comment was made about the project’s lack of comprehensive information (97 mentions), stating ”Too little attention to the negative sides of this train connection. Consider financial risk, landscape impact, etc”. The statement indicates that individuals who do not utilize trains may be critical about the adverse consequences associated with the construction of a new rail line. This could also account for the negative assessment of the PVE in the aforementioned categories.

Young Adults

The quantitative research reveals that young adults have a lower tendency to perceive PVE consultation as being important. Additionally, they are more inclined to perceive this consultation as difficult and less inclined to accept the government’s conclusions. Nevertheless, although young adults had a low assessment of the face validity of PVE consultations, it is noteworthy that the majority of comments in the Lelylijn consultation were provided by middle-aged individuals (1,180 comments), while young adults contributed just 517 comments. This suggests that young individuals are less inclined to express their opinions regarding the consultation compared to middle-aged individuals. Another potential cause of this low number of feedback may be that young adults perceive this topic as less important, leading them to neglect providing detailed comments during the consultation.

Regarding the category of comments, the most expressed comment was about the confusing choice experiment (96 mentions), while the complaint about the confusing allocation of points received 33 mentions. These comments shed light on why this group found the consultation difficult to understand. Furthermore, there were multiple references to the lacking information regarding the follow-up of the consultation (57) and the project (57), which may also result in lack of acceptance of the government’s decisions.

Summary

Key Takeaways

This chapter provided answers to **research question 2**: *”What are the factors contributing to the negative assessment of PVE by these specific segments?”*.

- The content analysis identified 23 distinct categories of criticism from the Lelylijn PVE consultation, grouped into five dimensions of face validity: consultation feasibility (34,7%), completeness of information provided (25,9%), concerns about acceptance (24,3%), transparency (10,8%), and challenges with clarity (4,2%).
- Despite generally giving higher overall consultation ratings, **highly educated individuals and men were particularly vocal in raising concerns**. Similarly, groups of respondents who were more likely to evaluate the PVE consultations negatively also provided the most comments (e.g., infrequent users of train and Lelylijn).

6 Experts Perspectives Exploration

After identifying certain characteristics of PVE participants who consistently evaluate PVE consultation negatively, and examining the causes behind their dissatisfaction, it is crucial to resolve these negative evaluations. In order to tackle these issues, a discussion was conducted with experts in PVE. This chapter will address the third research question, which is *"What strategies can be implemented to improve the evaluation of PVE's face validity for segments that consistently express negative views towards it?"*.

The selection of a group discussion with specialists in PVE as a strategy was based on its potential for mutual advantages. Sharing the results of the previous phase of this study would enhance the understanding of PVE practitioners, especially in the field of transportation, given that it is a relatively new approach. On the other hand, the strategies outlined in the forum have the potential to be more efficient because they involve individuals who have actively participated in the development and evaluation of PVE. Due to their expertise in this field, the specialists were able to promptly comprehend the findings and analyze how they may deviate or align with their existing knowledge. This would help to ensure that the study does not yield unnecessary solutions.

The group discussion comprised nine specialists in PVE consultation design and research. These experts were selected based on their extensive experience and significant contributions to the field. Among them were university and independent researchers who have developed PVE, published numerous papers on PVE methodologies and transport appraisal, and who work with governments in policymaking.

The forum began by presenting the quantitative and qualitative findings from Chapters 4 and 5. The initial presentation was followed by a Q&A session to clarify any points. Subsequently, the central question posed to the audience was **"How to address the negative evaluation given by certain groups?"**. The discussion was conducted in an open format, allowing participants the freedom to express their opinions without any restrictions. The primary topics discussed were as follows:

1. **The provision of additional information**

   The availability of sufficient information were critical in consultation method like PVE. The findings from the quantitative analysis in chapter 4 showed that roughly 70% of the participants think that they learned about the choices that the government have to make, while 10% think otherwise. This fact is reflected in the content analysis results in chapter 5, where there were significant numbers of respondents expressed that they find the consultation lacks of information (340 mentions), from about the project itself, its impacts, what experts think about it, and the follow up of the consultation. Given these findings, there are several thoughts expressed in the forum.

   The group conversation was started with a suggestion made by a PVE respondent, "Hyperlink to background or research information." While this could be a solution in providing more information, the use of hyperlink was actually not recommended as "you lose people from the consultation as they get to the hyperlink and they go to somewhere else". Another option raised to tackle that was to use a button that links to the methodological background instead. However, there is a bit of skepticism about providing more information in the consultation, in the group, as "giving more information, even if you have a perfect text, may actually sparks something that makes people more doubtful about anything there is". This actually relates to the study by Coulter and Ellins (2007) about providing too much information would generate more questions. Moreover, past experience shared by an expert in the group "we tried to give information about the governmental thing like how the decisions are going to be made, because people will just think that, okay, this is just a weird governmental world, I do not want to know about this". There are some values in providing more background information, but there will always be some risks. Eventually, the consensus seems to be that it is necessary to add more information, but it needs to be ensured

that any provided information should clarify more than it complicates.

2. **Having concise and extensive version**

   While the overall ratings given to all the 3 PVE consultations were good, in terms of face validity evaluation, about 9,7% of the respondents think that the consultation was difficult too understand. This fact also resonates with the results from content analysis, with feasibility dimension, which is how easy one answer questions or choose options, being the most commented (724 mentions) reason why participants did not like the consultation. This dimensions consists of comments ranging from the confusing type of questions, the lengthy nature of the consultation, up to the consultation being too complex. These findings also sparked some discussion in the forum.

   A comment given by a respondent in the PVE consultation was to have a concise and extensive version of PVE, so that people who are comfortable with the more simple version can stay with it, while people who needs more information can access the more nuanced version instead. A first thought raised from the panel was to reflect all of the feedback as ”it is quite risky, since all of these things are actually going to be beneficial to certain groups of (PVE) negative thinkers. Then, how would that affect the people who are now (already) think that the consultation is positive?”. The aforementioned comment was valid as currently the majority of the respondents have positive evaluation consistently across the consultation, something that seen as rare and ”not a normal distribution in other PVE consultation”. Indeed there needs to be consideration of downsides impacted by any solutions implemented for better PVE consultations. The consensus seems to be providing personalized information provided in the consultation, while maintaining simplicity where possible, are favored approaches. This personalization takes understanding ”who needs whats, and then in the beginning the participants can answer these questions”.

3. **Randomization of Options**

   The statement ”If many people participate in this consultation, it will increase my acceptance on the final decisions on this topic” received lowest average rates (64,8%) compared to other face validity statements. This finding also reflects in numerous comments about ”Acceptance” and ”Transparency” dimension. An example of comment from PVE respondent that could explain this was the order of the options in the consultation did not seem natural, ”option 1 will usually have the most points in your analysis because anchoring takes place here”. This finding adds new insight to the group as all these times, the order of the options are actually always randomized in each consultation, however the participants are not aware of this fact.

   A participant in the discussion group thought that it can be useful to put a button to provide information that the options are actually indeed randomized. However, a challenge raised that ”it is already a puzzle, how to put all of this information in there”. This can be addressed by only providing these additional information on the computer-version of the consultation. However, another concern aroused in the group that another question would be aroused on how some people have more access to information than other groups.

4. **Increasing Legitimacy and Trust**

   Results presented in the face validity evaluation in Chapter 4 showed that there is 15,9% of respondents think that the consultation is steering into a certain direction. This fact may relate to the numerous comments on ”Acceptance” and ”Transparency” dimension in the previous content analysis.

   The first solution aroused by the panel was to explain why it is steering or directive. This is due to the fact that ”PVE is steering by design. We limit the number of options and solutions because it is the essence of the PVE. So, perhaps, we should better explain that, why we do that and why we have certain options. So that the options are explained better”. A response to this idea was to ”think creatively about where we do explain in animation, instruction video, that we limit the options. We only presents options that are relevant for the policymakers”. Another response was to put an extra option like ”are you missing an option here?” which if a respondent clicks on it, there will be some explanation on why the sets of options were chosen. To quote the additional explanation, ”people can also give their own idea without effecting restrictions, so they still have

to make the same choices". As consequences, the PVE designers have to think about what questions the respondents will probably have in mind, while going through the process.

Another solution offered regarding this steering nature of PVE, was for the legitimate perspective that are not part of the choice design. To quote, "like *"I don't believe in climate change, but you know the choices tasks ask me to choose measures for fighting climate change".* I think we are aware that in many consultations we ask that kind of question like *"okay, what is your attitude towards this problem? And if not, would you like to still continue?"* but, perhaps we should do that more structurally, like, in the design phase, we should take all the legitimate perspective that are not part of the solutions or the choice universe that we present". A response aroused to that idea was that the perspective of those who might oppositions were actually wanted, as "someone who does not believe in climate change, they are still gonna be affected anyways by the windmills in the neighborhood. So you really want to hear what they think about them".

Another idea came up to increase the legitimacy of a method like PVE was to explore the effectiveness of "name-dropping" trustworthy institution in the consultation. To quote, "like we work with people from the municipality, we work with people from this organization and they chose these options for these reasons". Instead of just really explaining all the steps we take methodologically, just use other trusted or known actors to increase our legitimacy". A response to this idea was to think from the perspective of the organizations that are going to be mentioned in the consultation, as if they would be okay if citizens coming up to them asking if they really agree with the consultation.

Summary

Key Takeaways

This chapter addressed the **third research question**, focusing on strategies to improve the evaluation of PVE's face validity for segments that consistently express negative views. The group discussion with PVE experts involved sharing the quantitative and qualitative analysis results from the previous stages of this study, followed by a Q&A session and a discussion of potential strategies. The key points and strategies discussed were as follows:

- Provision of Additional Information: Enhance clarity about the process and the project while ensuring that the information provided does not overwhelm participants or create unnecessary doubt.
- Concise and Extensive Versions: Offer both concise and extensive versions of the consultation to cater to different participant preferences.
- Transparency in Randomization: Clearly communicate the randomization of options within the consultation to enhance transparency.
- Explaining the Steering Nature of PVE: Consider all legitimate perspectives, even those outside the presented options, and explore potential collaborations with trustworthy institutions to boost participant confidence.
- Expert Recommendations: The experts' suggestions are aimed at refining PVE consultation designs to make them more effective and better received by participants who may currently express negative views. These strategies emphasize clarity, inclusiveness, and transparency to improve the overall face validity of PVE consultations.

7 Discussion and Recommendation

This study aimed to gain insights into groups of people who (in)consistently evaluate PVE consultation in a negative or positive manner, specifically in transportation fields. This chapter provides discussion of the findings obtained from this study. The limitations and implications of this research are raised in this discussion. The beginning of the discussion focuses on the current state of knowledge of face validity in public participation methods, with regard to the first sub-question. Subsequently, there is an examination of the outcomes of data analysis concerning the relationship of respondents' characteristics with their assessment of PVE consultation. This includes an exploration of specific segments that emerge from the investigation. Next, discussion points concerning the (in)consistent assessment of different PVE consultations are presented. Next, the presentation of experts' opinions is discussed. This chapter concludes with a perspective on the connection between this study and broad welfare criteria, as well as a summary of the recommendations.

The first topic of discussion relates to the current state of face validity in the public participation approach. As stated in Chapter 3, validity is a comprehensive term. Validity encompasses various categories, one of which is face validity. Numerous studies have used face validity to assess the validity of an instrument in various sectors such as health, economics, and transportation. The rationale for this is that face validity is the most straightforward and quickest method for assessing validity (Bannigan & Watson, 2009; Patel & Desai, 2020). These investigations mostly aimed to assess the instrument's clarity, relevance, comprehensiveness, suitability, and understandability, even for general public. In addition, when it comes to public participation instruments, face validity is also used to determine if the instrument is suitable and effective in accurately capturing the opinions of respondents regarding policy options. Thus, it is important for a public participation method like PVE to be perceived as valid by its respondents. However, several studies have indicated that face validity alone is inadequate. Bannigan and Watson (2009) emphasized the importance of employing multiple validation procedures when assessing a measure, rather than depending solely on face validity. It is common practice to assess an instrument for both its face validity and content validity simultaneously (Connell et al., 2018; Dalawi et al., 2023; Hadi et al., 2023; Mason et al., 2020; Morton et al., 2023). Face validity alone does not offer a conclusive assessment of the overall validity of a PVE consultation. Therefore, further research may be essential to investigate other types of validity in PVE.

The next discussion addresses the second sub-question concerning the association between respondents' characteristics and their evaluation of the face validity of PVE. To answer this question, it is important to consider whether to weight the dataset, particularly if there are significant differences between the characteristics of the respondents and those of the general population in the Netherlands. Initially, the analysis was conducted using the actual (non-weighted) dataset. However, the results from this non-weighted dataset were not meaningful, as there were few significant associations between respondents' characteristics and their evaluation of face validity statements. Additionally, Tuit (2022) also employed weighted datasets to draw conclusions for populations, reinforcing the appropriateness of this approach. Therefore, this study utilized weighted data for all statistical analyses. However, potential concerns arise regarding the factor weighting of the dataset. Different datasets may have varying sets of respondents' characteristics, sometimes differing in levels from the available benchmarks. Consequently, some simplifications were made during data analysis, which may result in slight differences in outcomes. It is recommended that future studies maintain consistent levels and characteristics across consultations and align them with current demographic groups to minimize discrepancies. This alignment will help ensure that the results are more representative and the conclusions more robust.

The next discussion focuses on the dataset used for the analysis to answer the second sub-question. During the data preparation step, the dataset was filtered to exclude responses of "I don't know." In the actual PVE consultation, several options in both face validity statements and questions regarding respondents' demographics included the "I don't know" answer. This filtering was done to ensure consistency across different statistical analyses and various PVE dataset analyses. As a result, the

number of data points used for analysis have changed, although the dataset still included a substantial number of respondents (approximately 4.000 respondents for each dataset). In the descriptive analysis results presented in Chapter 4.2.4, it was observed that the overall ratings for all three consultations were quite high (7,42, 7,72, and 7,55 for the Lelylijn, Oude Lijn, and Mobility Vision consultations, respectively). When these results were presented to PVE experts, it was noted that such high ratings are relatively rare compared to past PVE consultations. One possible explanation is that these three consultations predominantly presented ”good news” to respondents, such as the introduction of more train stations, new train lines, and increased accessibility. In contrast, other PVE consultations also discussed the drawbacks of the projects being evaluated. However, another potential reason could be related to the exclusion of responses from respondents who chose the ”I don’t know” option in any of the questions. Further research is needed to determine whether this approach significantly affects the overall ratings of PVE consultations.

The rare trend of high overall ratings across the consultations, compared to past PVE consultations, could also be influenced by the existence of potential respondents who might have rated the PVE consultation lower but decided to drop out during the process. In this study, the dropped-out respondents were not included in the analysis. Hence, there is a possibility that some respondents found the consultation too complicated or nonsensical, to the extent that they disliked the process and chose to quit midway. Currently, within the three PVE consultations analyzed in this study, there were no follow-up questions addressing why participants decided to quit the consultation. This omission leaves a gap in understanding the rate of dropout and the reasons behind respondent dropout, which could be crucial for interpreting the overall high ratings. To address this issue, future PVE designs could incorporate a final multiple-choice question for respondents who choose to quit the consultations. This question could explore whether they found the consultation too complex, too lengthy, irrelevant to their interests, or other reasons. Additionally, recording the specific step at which respondents quit would provide valuable insights into the stages of the consultation that might need improvement.

The next discussion focuses on the evaluation of face validity in the PVE consultations. According to the analysis results in Chapter 4, 15,9% of respondents in the Lelylijn consultation agreed with the statement that the consultation was influencing their choices, whereas 37,4% of respondents were neutral. In comparison, Golan (2023) found that 81% of participants in the Tel Aviv transport PVE agreed with the statement, ”The research was objective and did not steer my choices in a certain direction.” Despite the different formulations of the statements, both studies show a relatively low perception of the consultations being steering and not objective; however, the significant proportion of neutral responses in the Lelylijn consultation indicates that many respondents may have reservations or uncertainties about the consultation’s objectivity. Another possibility is that the nature of the consultations influenced these perceptions, as the Lelylijn consultation primarily presented options assuming that the Lelylijn will be built, whereas the Tel Aviv PVE included various ”good news” scenarios for public transport. This difference is echoed in some respondents’ comments in the Lelylijn consultation, stating that ”the consultation is steering as it already assumed the Lelylijn will be built without asking if we want it in the first place.” Experts interviewed in Chapter 6 confirmed that PVE consultations are indeed steering by design due to the limited options and solutions provided, suggesting that this steering nature should be explicitly communicated to respondents to enhance transparency. It would be intriguing to observe whether future PVEs with more favourable natures, similar to the Tel Aviv PVE instance, still lead to a significant proportion of participants considering the consultation as steering or neutral, particularly if the face validity statement is constructed positively.

Another negatively formulated statement, “this consultation is difficult to understand,” existed in the Oude Lijn and Mobility Vision datasets. The presence of this statement complements the check of the feasibility dimension of face validity, thus fulfilling the recommendation from Tuit (2022). The results in Chapter 4 showed that almost 10% of the respondents in both datasets found the consultation difficult to understand. However, the content analysis in Chapter 5, which was conducted on the Lelylijn consultation, revealed that many highly educated individuals expressed concerns that “many people think that this consultation is difficult” or voiced concerns about “people who might be digitally illiterate might find it difficult to participate in this consultation.” This raises the question of whether the consultation is genuinely difficult or if it is an assumption by certain groups about other groups perceived to be less educated or less digitally literate. To address this ambiguity, it would be interesting for future studies to examine the demographics of people who actually answered that the consultation is difficult and doing

content analysis for Oude Lijn and Mobility Vision dataset. Analyzing different demographic factors such as education level, gender, and digital literacy could provide insights into whether certain groups are disproportionately affected by the complexity of the consultation.

Regarding digital literacy, some respondents in the Oude Lijn and Mobility Vision consultations in this study were drawn from a panel of respondents. Using a panel of respondents means that these individuals were pre-selected based on their ability to use an online web tool, as panels typically consist of people who frequently participate in online surveys. Therefore, there are no individuals in the panel who are entirely unfamiliar with how a PVE consultation should be conducted online. This pre-selection potentially introduces a bias, as panel participants are likely to have higher digital literacy. A hypothesis is that panel respondents might report higher face validity compared to those who are less digitally literate. This could result in an inflated assessment of face validity. Nevertheless, there are also respondents from the open consultation, meaning that these individuals voluntarily joined the consultation through other media. To ensure inclusiveness and a broad distribution of the consultation, it might be beneficial to send invitations to join the PVE consultation through traditional mail (e.g., post.nl) or link it to citizens' IDs to control for duplicate responses. However, this approach must be carefully managed to maintain the anonymity and ethical considerations of the respondents.

Moving on to the relationship between respondents' characteristics and their evaluation of PVE consultations, the results in Chapter 4 indicate that some characteristics were found to be consistent, inconsistent, or even insignificant across consultations. One consistent finding is that young adults tend to rate PVE consultations highly overall. However, multinomial logistic regression results show that young adults provide less positive evaluations compared to elderly people. This finding resonates with Tuit (2022) study, where young people rated all five face validity categories lower. A notable difference from Tuit (2022) study is observed in the proximity factor. In her study, people who lived close to a problem situation tended to rate face validity lower. This contrasts with the observations in the Lelylijn and Oude Lijn consultations, where people consistently evaluated face validity statements more positively. One potential reason for this discrepancy could be the definition of proximity. In this study, North Holland was defined as "proximate" for Lelylijn and South Holland as "proximate" for Oude Lijn, due to the larger geographic scope of these projects. This broader definition of proximity might differ significantly from Tuit (2022) study in the more concentrated Schiphol area. However, it might be important to consider not just where someone lives, but also how close their regular activities, such as going to school or work, are to the project in question. This approach could provide a more nuanced understanding of how proximity affects respondents' evaluations of face validity.

In the results analysis presented in Chapter 4, some respondents' characteristics were found to be insignificant or showed varying associations across different consultations. These characteristics include financial status, type of home, occupancy, and area of living. While it is possible that there is indeed no association between these characteristics and respondents' evaluations, an alternative hypothesis is that associations might emerge if the levels of each characteristic are presented in more detail. For simplification, this study categorized characteristics into three levels (e.g., low, medium, high). However, it might be insightful to explore the results with more granular levels of these attributes. For instance, defining financial status with a range of monthly incomes rather than broad categories of enough, not enough, and more than enough money. Such detailed categorization could potentially reveal associations that are obscured by the broader groupings used in this study. This approach could provide a more nuanced understanding of how these characteristics influence respondents' evaluations of face validity in PVE consultations.

Finally, the last discussion touches on the results of the content analysis of the negative feedback from the Lelylijn consultation. This study focused solely on the negative feedback for the Lelylijn consultation due to limited resources. Nevertheless, the number of responses analyzed was sufficiently high, ensuring a substantial capture of feedback. Given that the Lelylijn consultation was conducted first, followed by the Oude Lijn and Mobility Vision consultations in subsequent years, notable differences were observed among the three consultations. These differences are expected to have addressed some of the negative comments from the Lelylijn consultation. Therefore, it would be valuable to conduct similar content analyses on the other datasets to determine if different outcomes arise from the negative evaluations. Additionally, content analysis is inherently qualitative and subjective. The categorization and interpretation of comments are subject to the researcher's perspective, particularly since the com-

ments were English translations from the original Dutch. To ensure the consistency and robustness of the results, further content analysis might benefit from peer review (Barbour, 2001).

Reflection on PVE relation with Broad Welfare (BW) Criteria

The relevance of this research in supporting the development of forward-thinking and inclusive transport designs, which are consistent with the broad welfare *("brede welvaart")* perspective, or welfare beyond GDP, was emphasized in multiple instances in the previous sections. The PVE methodology has been found to be well-suited for the evaluation of public policy based on broad welfare in previous studies. This is due to the fact that it allows different demographic groups to share their subjective experiences of prosperity, thereby ensuring that all essential values, preferences, and concerns are considered during policy deliberations (Rijkwaterstaat, 2022). In addition to assessing the face validity of the PVE method in this study, it is interesting to also reflect whether face validity evaluations and PVE consultations correspond with the Broad Welfare approach.

TNO identifies six elements that comprise the concept of broad welfare, that can be used as criteria for evaluation of tools for a comprehensive analysis of potential synergies with the broad welfare (See Appendix I). The first one is **"Time periods,"** which assesses if a tool or technique explicitly incorporates time periods and accommodates multiple time frames. Although the face validity statements in PVE consultations did not explicitly aim to evaluate the temporal nature of PVE as a tool, the flexibility of PVE allows for customization to incorporate different factors, including time considerations, if they are considered relevant in a specific policy context. For instance, the third dataset utilized in this study is the Mobility Vision project, the findings of which will aid governments in formulating transportation plans for the Netherlands from the present until 2050. Therefore, PVE meets this element.

The second element is **"Area Based"**, which evaluates if the instrument or method explicitly considers and accommodates different geographies. During PVE consultations, the specific locations of the projects being studied are clearly disclosed. For example, in the first two datasets of this study, Lelylijn and Oude Lijn, the locations are explicitly mentioned. These projects are concentrated in the North and South Netherlands, respectively, thus focusing on region-specific impacts. Certain sections of the Oude Lijn consultation explicitly enquired about individual stations, such as Leiden station. Furthermore, PVE facilitates the collection of viewpoints from both local and distant residents, enabling a comprehensive examination of diverse and even potentially comparable perspectives on the project. In addition, PVE can be employed at the local level (Tuit, 2022) and even at the national level, as demonstrated in projects such as the Mobility Vision project. The flexibility in PVE's design allows it to adhere with the "Area Based" requirements.

The third element is **"Participation and Co-creation,"** which checks whether a tool or approach makes the level of participation and co-creation explicit and allows for various levels of participation and co-creation. The nature of PVE, which puts citizens "in the seat of the government" by allowing them to provide advice on government choices, indicates that PVE enables participation and co-creation in the policy evaluation process. Additionally, some stakeholders can even be part of the PVE design process (Populytics, 2021). Furthermore, this element is reflected in the face validity statements: "This method should be used more often to involve residents in government policies" and "If the government involves residents in choices on a large scale through this consultation, then the final decisions about the Lely Line will be more acceptable to me." Therefore, indeed PVE aligns well with the participation element (Spruit & Mouter, 2020). However, whether the co-creation element is achieved or not by PVE, solely depends on how the decision-maker will actually utilize the PVE results for their decision-making.

The fourth element of Broad Welfare building block, **"Target groups and distribution effects,"** assesses whether a tool explicitly considers different population groups and analyzes the distribution of benefits and burdens among them. PVE meets this element by collecting socio-demographic data during the consultation phase, enabling a detailed segmentation of respondents. This allows for a thorough analysis of how different groups think about the project. By identifying the specific concerns and benefits experienced by various demographic groups, PVE provides insights into how impacts are distributed among different population segments. These findings illustrate distribution effects, where the same project yields varying impacts—both positive and negative—on different groups. By explicitly addressing these distribution effects, PVE ensures that equity considerations are integral to the evaluation process, making it a robust tool for assessing the broad welfare impacts of transport projects.

The fifth element, **"Integral Approach & Systems Perspective,"** evaluates if a tool takes into account the interconnections between different factors and topics, and contains a comprehensive, systems-oriented viewpoint. More precisely, it assesses whether the tool offers a means to incorporate various separate entities within the government (such as multiple ministries working on the same subject) or synchronize different levels (such as national and local levels collaborating). Although PVE does not directly offer a specific method for integrating several government silos or layers, it frequently requires the participation of multiple ministries and specialists in the design phase. In addition, PVE enables citizens to assess portfolios of projects or policies that may encompass various government departments or levels, promoting a comprehensive approach to policy evaluation. This process can illuminate the synergies or conflicts that may arise between initiatives or interests from different government entitities. By collecting citizen preferences on a range of policy options in a single exercise, PVE provides a unified source of public input that can be valuable for various government departments and levels, thus indirectly supporting an integral and systems perspective.

The final element, **"Synergies and Trade-offs,"** assesses if a tool encourages the identification of synergies and clearly articulates trade-offs, particularly when it comes to allocating scarce resources. This element is met in PVE when respondents have the option to choose many projects within a specific budget or point limit. This allows respondents to weigh the trade-offs and potential synergies among the initiatives. This element is also associated with one of the face validity statements, "Through my participation in this consultation, I have gained knowledge regarding the decisions that the government must make regarding this matter". Nevertheless, it is important to acknowledge that the positive and negative synergies of the initiatives themselves are not within the scope of PVE (Bahamonde-Birke & Mouter, 2024). Therefore, another approach or tools might be necessary to capture the potential synergies among the initiatives.

While PVE is utilized as a transportation planning tool within the broad welfare perspective, it is important to recognize that this study cannot be directly compared to evaluations of tools using BW criteria. This study focuses on the face validity of PVE, specifically how users or respondents perceive the tool in terms of its effectiveness in measuring what it intends to measure. In contrast, BW criteria are employed by experts or organizations to assess various tools for comprehensive analysis of potential sinergies with the broad welfare. Nevertheless, there are notable overlaps and connections between this study and BW criteria, suggesting that PVE is indeed aligned with the principles of broad welfare. Further research could be valuable in exploring this relationship and PVE's contribution to the field of broad welfare research more comprehensively.

Summary of Recommendations

This chapter synthesized the findings from this study and provided actionable strategies for improving future PVE consultations.

Recommendations for future PVE:

- Provide balanced information to clarify processes and project details, ensuring it is accessible without overwhelming participants or creating unnecessary doubt.
- Develop both concise and extensive versions of the consultation to accommodate participants with varying levels of engagement and information needs.
- Clearly communicate the randomization of options and the inherent steering nature of PVE.
- Collaborate with trusted institutions and explicitly explain the rationale behind the options provided to enhance the legitimacy of the consultation process.
- Recruit participants more ”formally” through traditional post or linking the consultation with citizens’ ID.

Recommendations for future studies:

- Explore additional types of validity for PVE, such as content validity, to strengthen the overall robustness of PVE as a tool for public participation in transportation planning.
- Refine Segmentation and Analysis: More granular categorizations of participant characteristics, and ensure that the levels across consultations are comparable for more significant comparisons. Alternatively, analysis without exempting ”I don’t know” answer could also be performed.
- Investigate the reasons behind participant dropout to improve the design and inclusiveness of future PVE, by adding follow up question before leaving the consultation.
- Investigate the consistency of results if the statements are re-formulated (e.g., consultation is not steering and objective).
- Explore more about PVE and its relation and contribution to broad welfare.
- Explore further the group of people who evaluates the particular face validity statement (e.g., education level of people who think the consultation is difficult).
- Perform follow up content analysis for the disliked aspects in Oude Lijn and Mobility Vision consultations, complemented with peer reviews.

8 Conclusions

The Participatory Value Evaluation (PVE) method is a relatively new approach that serves as both a public participation tool and an evaluation method in transportation policy-making. Despite its increasing use, empirical research into the validity of PVE and its perception among different population segments remains limited. Chapters 1 and 3 highlight the significance of (face) validity of public involvement methods, such as PVE, in order to ensure that the outcomes are credible and can effectively assist transportation policy-maker. One of the types of validity is face validity, which evaluates whether a tool appears to measure the intended outcome, particularly in the eyes of its users, in this instance, the participants of the PVE consultation. This study has yielded valuable insights into how different demographic groups evaluated the face validity of PVE consultations in transportation and accessibility initiatives. This chapter presents the conclusion for each research question in this study, along with the comparison with the hypotheses defined in the beginning.

1. **To what extent are there distinct segments in the population who (in)consistently evaluate the face validity of Participatory Value Evaluation (PVE) in transportation planning?**

   This question will be addressed by answering three sub-questions as outlined below.

   1.1. **What is the current state of knowledge, in literature, regarding face validity in public participation approaches, particularly within the realm of transportation planning?**

   This thesis has undertaken a literature review to identify the current state of knowledge regarding the face validity of citizens participation methods. The literature review demonstrates that face validity is a simple and quick approach for evaluating the validity of an instrument. Face validity is commonly employed to assess the relevance, clarity, comprehensiveness, suitability, and understandability of an instrument, particularly for general people. Nevertheless, it is acknowledged that depending exclusively on face validity might be inadequate for evaluating the whole validity of a tools, like PVE.

   Face validity has been utilized in a wide range of instruments, spanning from health questionnaires to surveys related to autonomous vehicles. A framework has been created to evaluate the face validity of public participation tools that engage citizens in expressing their thoughts on government choices during PVE consultations. This framework includes statements regarding the clarity, unambiguity, relevance, readability, and completeness of the PVE consultation, as seen in the Schiphol Social Council study. In the Tel Aviv public transport PVE study, face validity was assessed by examining criteria such as completeness, acceptance, transparency and legitimacy, relevance, and readability.

   Currently, a more comprehensive framework for evaluating face validity in PVE is being prepared for publication. This framework consists of nine categories: clarity, completeness, feasibility, relevance, transparency, acceptability, legibility, sensitivity, and familiarity. Despite these developments, there are still limited studies on the face validity of public participation methods in transportation decision-making, especially when it comes to the population segments that consistently assess PVE's face validity in a certain way, the reasons behind it, and how to address the negative perceptions.

   1.2. **What are the characteristics of the distinct segments in PVE consultations, and how do these characteristics relate to their face validity evaluations?**

   The study on segments in PVE evaluation encompassed various data analyses. The initial quantitative analysis involves calculating descriptive statistics for the overall consultation ratings across the three PVE consultations. The results indicate a fairly positive outcome, with average scores of 7,42, 7,72, and 7,55 for the Lelylijn, Oude Lijn, and Mobility Vision

consultations, respectively. These high scores are considered rare, as PVE experts have noted that they are considerably higher than previous PVE consultations.

Crosstabulation analyses were performed to examine the relationship between respondents' attributes and their ratings of the overall consultation. The findings indicate that the **following groups of people tend to rate the consultation higher**: young adults, men, highly educated people, home renters, unemployed individuals, frequent train users, people who (plan to) use Lelylijn/Oude Lijn more frequently, users of the train stations being evaluated, urban dwellers, and people who rarely have accessibility issues. The results confirm the initial hypothesis made at the beginning of the study.

In order to determine how respondents responded with the seven face validity statements in the four face validity dimensions across the three consultations, descriptive analyses were performed. Regarding the dimensions of **transparency and feasibility**, the findings indicate that, on average, 83,9% of respondents think that the consultation is a fair and honest investigation. This suggests a high level of confidence in the integrity of the consultation. This statement has the highest average agreement rate among the seven face validity statements. However, 15,9% of respondents in the Lelylijn consultation thought the consultation is steering. On average, 9,7% of respondents in the Oude Lijn and Mobility Vision consultations found the consultation difficult to understand.

Regarding the **relevance dimension**, 93% of respondents in Lelylijn consultation considered the topic to be important to them, despite the fact that a majority of respondents in Lelylijn reside far from the project. Only 72,8% of respondents in Oude Lijn believe that the topic is important for them to express their views on. Regarding the statement "by participating in this consultation, I learned about the choices the government must make on this topic," approximately 70,3% of respondents in Lelylijn and Oude Lijn agreed with it, whereas 9% of respondents in Lelylijn and Mobility Vision disagreed.

The final aspect of face validity examined in this study related to the level of **acceptance**. Approximately 83% of the respondents expressed agreement with the statement "this method should be used more often to involve residents in government policies." This indicates that the respondents are receptive to and supportive of PVE as a means of public participation. However, when it comes to the final decisions themselves, only an average of 64,8% of respondents agree with the statement "if many people participate in this consultation, the final decisions on this topic are more acceptable to me." This suggests that a smaller proportion of respondents believe that their participation in this method will increase their acceptance of the government's final decisions. Alternatively, there is a possibility that the respondents already have high acceptance on the final decisions, so that participating in the consultation would not increase their acceptance. Finally, a mere 67,4% of participants agreed with the statement that if the government includes residents more frequently in deliberating about such choices, their confidence in the government's actions will increase. This suggests that amost of respondents believed that engaging in PVE activities leads to an increase in their trust in the government.

In order to determine if there are distinct groups of individuals who evaluate face validity statements and overall ratings in PVE in certain ways based on their traits, Latent Class Cluster Analysis (LCCA) were performed. The LCCA analysis yielded four distinct clusters at each PVE consultation. The largest cluster usually consists of individuals who highly evaluate PVE. The second largest clusters in each consultation consist primarily of positive evaluators, however certain face validity statements are evaluated with neutrality. During the Lelylijn consultation, the third and fourth clusters are both characterized as negative evaluators. The main difference between the two is that the third cluster consists of frequent train users, while the smallest group consists of non-train users. In the Oude Lijn dataset, the third largest cluster consists of critical evaluators, whereas the smallest cluster consists of neutral evaluators who are non-train users. In the mobility vision consultation, the third largest cluster consists of neutral evaluators, while the smallest cluster consists of negative evaluators. The results of the LCCA exhibit a comparable pattern evaluation to that of crosstabulation analysis. It reveals that individuals belonging to positive groups are typically young adults

with a medium to high level of education. They also reside in close proximity to the project and are infrequent users of the train and the line. The members of the negative evaluators group are often young to middle-aged individuals, with an equal likelihood of having low or high levels of education. They tend to reside distant from the projects and are infrequent users of the train and the line.

1.3. How consistent are these face validity evaluations across different PVE experiments in transportation planning?

Crosstabulation Analyses

While LCCA examined the clusters of respondents by considering all the statements and characteristics of all respondents together, Crosstabulation analyses were conducted to explore the relationship between respondents' characteristics and their evaluation of the 7 face validity statements in the PVE consultations. These analyses discovered several consistent, inconsistent, and insignificant associations between the respondents' attributes and their assessment of face validity in PVE. The findings indicate that young individuals are more inclined to believe that they learned about the government's decisions by actively engaging in PVE consultation. Males are more likely to think that their level of trust in the government rises as a result of PVE participation. However, they are less inclined to believe that they learned about the government's decisions. Furthermore, those with a higher level of education are more likely to perceive the topic as important and find the consultation process effortless. Nevertheless, their willingness to engage in PVE consultation is unlikely to result in an increase of acceptance and trust in the government's decisions. While these findings align with the findings of the earlier crosstabulation analysis about the overall consultation rating, it is clear from the results that while men and highly educated individuals generally like the PVE consultation, they are critical of learning and of growing in acceptance and trust in government decisions.

Another consistent evaluation is in line with the findings of cross-tabulation of overall ratings. For instance, individuals who reside in close proximity to the project, people who have a higher rating tend to perceive the consultation as both honest and important. Likewise, regular train commuters tend to believe that the topic is important, that the investigation is honest, that they gained knowledge from the discussion, and it should be utilized more often. Finally, the more the frequency with which an individual intends to utilize the Lelylijn/Oude Lijn, the stronger their perception of the topic's importance, the honesty of the consultation, and the necessity for its further usage.

Although there were numerous associations that were consistent, there were also several associations that were inconsistent or insignificant across the three consultations. The type of residence a person lives in and their occupation appear to be associated inconsistently across the consultation. Individuals residing or working in close proximity to the examined stations are less inclined to consider Oude Lijn PVE as an important topic to express their viewpoint on, but this attribute does not exhibit a substantial relationship with other statements. Furthermore, attributes such as previous utilization of the stations, financial status, and frequency of encountering issues with accessibility are only significantly correlated with three statements during a single consultation.

Multionomial logistic regression

While crosstabulation analysis reveals the relationship between respondents' attributes and their face validity assessment, it does not control for the influence of other traits. Therefore, multinomial logistic regression was conducted to examine the impact of each attributes of respondents on their evaluation of face validity statement. Additionally, the analyses aimed to identify if similar patterns appeared across different analyses. The analyses, similar to crosstabulation analysis, produced results that showed consistent, inconsistent, and insignificant relationships between the characteristics of the respondents and their evaluation of face validity in PVE.

The groups that were previously shown to have high rates appear to receive slightly nega-

tive evaluations in the results of the multinomial logistic regression. Young individuals often assign higher scores to the overall consultation, but they are less inclined to perceive the consultation as honest, easy, important topic, and effective in increasing their acceptance of government decisions. Thus, it appears that young individuals are more discerning in their assessments of face validity. While men tend to provide better overall consultation ratings, they are less inclined to perceive the consultation as honest, easy, and to be used more. They are unlikely to perceive that they acquired knowledge about the government's decisions through the consultation. Contrary to the previous crosstabulation result, it is less probable for them to believe that engaging in PVE activities enhances their trust in the government. Following that, Individuals with a higher level of education are more inclined to perceive the consultation as easy to understand, mirroring the findings of the crosstabulation analysis. However, while they are more inclined to give a higher rating to the consultation, they are less likely to believe that they have gained knowledge about the government's decisions. Additionally, they are less likely to think that the consultation should be utilized more frequently and that engaging in PVE activities increases their trust in the government. As with the crosstabulation results, the topic is perceived as important by those who reside closer to the project. Finally, individuals who often utilize trains are more prone to believe that the consultation process is honest and are likely to perceive that engaging in PVE consultation enhances their trust in the government.

The other features of the respondents have a inconsistent or insignificant impact on how they evaluate the face validity statements. Regarding employment, while unemployed individuals perceived this consultation as difficult to understand, they considered it to be honest. Similarly, the type of home in which someone resides also has varying effects on their face validity evaluation. Finally, individuals who encountered less accessibility issues tend to exhibit a more favorable attitude while evaluating the face validity assertions in the Mobility Vision consultation.

2. **What are the factors contributing to the negative assessment of PVE by these specific segments?**

To determine the factors that contributed to the negative evaluations of PVE, a content analysis was conducted on the responses provided in the Lelylijn consultation. The study revealed that the feasibility dimension received the highest number of comments (724), followed by completeness (540), acceptability (507), transparency (226), and clarity (87). There is a consistent proportion and order among type of comments from different categories. Therefore, each group exhibits identical types of comments within the overall dataset.

The prior crosstabulation findings indicate that individuals with higher levels of education and men are more inclined to give better ratings for the overall consultation. However, it appears that same groups are also more outspoken in expressing their dissatisfaction with the consultation, since around 75% of the comments originate from them. On the other side, the other top commentators were usually those who were not potential users and tended to give the consultation a low rating. The comments expressed by the individuals who gave high ratings were consistent with their assessment of face validity. Both of these groups are unlikely to believe that they acquired knowledge about the government's choices through the process of consultation. This is evident in the large number of comments regarding the consultation's completeness, the majority of which focus on issues such as the project's lack of information, its current status, potential consequences that were overlooked, and the process by which the options were developed.

The most common comment in the entire dataset are related to feasibility dimension. The majority of these criticisms are to the sections on choice experiments and point allocations, which were deemed confusing by many individuals. Other feedback highlights the consultation's lengthy duration and complex nature. Furthermore, there are further remarks regarding the technical aspects and user experience during the consultation, such as the inability of individuals to immediately view the given points.

Regarding the acceptance component, there have been numerous remarks regarding the virtual nature of the consultation. As a result, some individuals have raised questions about the par-

ticipants involved in this survey. Due to the possibility that certain individuals affected by this research may lack digital literacy skills. However, there are also concerns about the consultation being dominated by individuals with specific interests. These factors contribute to their reluctance to accept the final decisions.

Regarding the acceptance of the PVE itself, there have been various criticisms regarding the formulation of the questions and the limited opportunity to express one's own opinion. While regarding the Transparency dimension, the majority of comments mostly focus on the perception that the consultation process is being steered or directive. Numerous complaints have been made about having to select an option that they may not like, about being asked about Lelylijn when they do not agree with Lelylijn, and other similar experiences. Regarding the dimension of clarity, numerous comments have been made regarding the lack of clarity in the questions.

3. **What strategies can be implemented to improve the evaluation of PVE's face validity for segments that consistently express negative views towards it?**

   A number of strategies can be used to address the negative assessment of PVE's face validity and PVE consultation in general . Firstly, as the term "feasibility" received the most criticisms from the respondents, it would be advisable for future PVE to have both a brief version, which is shorter and more succinct, and an extensive version, which encompasses all the necessary detailed information. This is related to the comments on the "completeness" component, that received numerous concerns concerning the adequacy of information provided regarding the project and consultation process. Enhancing the amount of information provided could enhance the likelihood of respondents understanding of the decisions made by governments. Moreover, it could increase the respondents' acceptance and trust in the government's decisions. However, the experts also expressed concern about the potential drawbacks of including much information, as an abundance of information could potentially generate additional inquiries. Hence, it is recommended to offer additional information, but still ensure that it clarifies rather than confuses.

   Another measure that can be taken to enhance feasibility dimension is to simplify the consultation process. Given the findings that suggest individuals with lower levels of education perceive the consultation as difficult, the PVE designer could incorporate buttons that, when hovered over, provide clarification on specific tasks, additional explanations of complex terms, or even a demonstration of how to allocate points. In order to enhance the respondents' understanding, it is advisable to provide additional visuals, maps, and visual representations to effectively describe the topic and options being discussed.

   In order to respond to the feedback on transparency and acceptability, there are a few actions that can be taken. Firstly, in order to acknowledge the significant number of respondents who believe that the consultation is steering, the future design of the PVE should explicitly state that the nature of PVE is indeed directive, as the options presented during the consultation were selected based on valid and compelling reasons. The PVE designers might consider mentioning the names of reputable institutes to enhance the credibility of the selected solutions. In order to solve the criticism regarding respondents' incapacity to express their own thoughts, it may be necessary to find a mechanism to incorporate the personal viewpoints of respondents that may not be covered by the available alternatives in the consultation.

Conclusions on Hypotheses

H1: Older individuals tend to evaluate PVE more positively compared to younger individuals.

The findings in this study indicate that younger adults are more likely to give higher overall PVE consultation ratings. However, multinomial logistic regression results show that young adults are less likely to perceive the consultation as honest, easy, important, and effective in increasing their acceptance of the government's decisions, compared to elderly individuals. Therefore, while the overall consultation ratings by younger adults diverge from the initial hypothesis, the analysis of several face validity statements supports the hypothesis to some extent. This suggests that while younger individuals may rate the consultation higher overall, older individuals may have a deeper appreciation for specific aspects of the PVE process.

H2: Women generally have a positive attitude towards PVE compared to men

The crosstabulation results on overall ratings show that men tend to rate the overall PVE consultations higher compared to women. However, multinomial logistic regression results indicate that men are less likely to perceive the consultation as honest, easy, and something that should be used more frequently. Additionally, men are less likely to think that they learned about the government's choices through this consultation and less likely to experience an increase in trust in the government as a result of participating in PVE. Therefore, similar to the findings on age, while the overall ratings do not align with the initial hypothesis, the results of the multinomial logistic regression suggest that women are more positive about several aspects of the face validity of PVE, thereby confirming the initial hypothesis to some extent.

H3: Highly educated respondents view PVE positively but with a more critical perspective.

The findings in this study show that individuals with higher education levels tend to rate the overall consultation higher compared to those with lower education levels. These findings are mostly reflected in the results of face validity evaluations, where highly educated individuals tend to think the consultation is easy. However, they are less likely to believe that they learned about the government's choices, that this consultation should be used more often, that it increases their acceptance of the final decisions, or that participating in PVE increases their trust in the government, compared to individuals with lower education level. Therefore, while the overall rating findings resonate with the hypothesis, the face validity evaluation results diverge from the hypothesis to some extent, indicating a more complex relationship between education level and perceptions of PVE.

H4: Frequent users of the train or specific project tend to have a more positive evaluation of PVE.

The findings of this study indicate that frequent train and studied line users are more likely to rate the overall consultation higher. These findings resonate with the respondents' face validity evaluations, where frequent users are more likely to evaluate the statements more positively in most cases compared to infrequent users. Therefore, the findings align with the initial hypothesis, suggesting that familiarity and regular use of the train positively influence the evaluation of PVE.

H5: Individuals living closer to the project are more likely to have a positive perception of PVE.

The findings in this research show that individuals who live closer to the project under scrutiny in the PVE consultation tend to give higher ratings to the overall consultation. This result aligns with the face validity statement evaluations in terms of perceived importance and honesty of the investigation, although it remains impartial in the rest of statements. Therefore, the results support the initial hypothesis, suggesting that proximity to the project positively influences individuals' perceptions of PVE.

Final Thoughts

This study has provided valuable insights into the (in)consistency of perceptions regarding face validity in PVE consultations across different demographic segments, addressing a significant gap in the literature as identified at the outset. The findings underscore the complexity of evaluating public participation tools like PVE, revealing that while overall ratings tend to be positive, nuanced differences exist in how specific aspects are perceived by various groups.

The research highlights the critical importance of tailoring PVE consultations to better accommodate the diverse needs of participants, particularly by addressing challenges related to feasibility, transparency, and the completeness of information. By enhancing these areas, future PVE consultations can improve their effectiveness and credibility, ultimately contributing to more inclusive and informed transportation policy-making. Furthermore, the reflection on PVE's alignment with broad welfare criteria reinforces its suitability as a tool for transportation planning within this broader perspective.

As PVE continues to evolve as a public participation tool, ongoing research and refinement are essential to ensure it remains responsive to the needs of all participants, fostering greater trust and engagement in the policy-making process. This study lays the groundwork for future exploration and underscores the potential for PVE to play a crucial role in shaping sustainable and equitable transportation solutions.

References

Aacharya, H. (2022). Content analysis. *EPRA International Journal of Research and Development (IJRD)*, *7*(4), 177–180. https://doi.org/10.36713/epra2016

Aini, M. F., Elias, M., Lamers, H., Shariah, U., Brooke, P., & Hafizul, H. M. (2017). Evaluating the usefulness and ease of use of participatory tools for forestry and livelihoods research in sarawak, malaysia. *Forests, Trees and Livelihoods*, *26*(1), 29–46. https://doi.org/10.1080/14728028.2016.1246213

Anastasi, A., & Urbina, S. (2007). *Psychological testing* (2nd). Pearson.

Ansari, W., & Andersson, E. (2011). Beyond value? measuring the costs and benefits of public participation. *Journal of Integrated Care*, *19*(6), 45–57.

Arnstein, S. R. (1969). A ladder of citizen participation. *Journal of the American Institute of Planners*, *35*(4), 216–224. https://doi.org/10.1080/01944366908977225

Avcıoğlu, O. (2023). *Public participation in local decision-making processes: The case of bodrum* [Master's Thesis]. Middle East Technical University. URL_TO_THE_THESIS_IF_AVAILABLE

Babalola, O., Goudge, J., Levin, J., Brown, C., & Griffiths, F. (2022). Assessing the utility of a quality-of-care assessment tool used in assessing comprehensive care services provided by community health workers in south africa. *Frontiers in Public Health*, *10*, 868252. https://doi.org/10.3389/fpubh.2022.868252

Bahamonde-Birke, F. J., & Mouter, N. (2024). About positive and negative synergies of social projects: Treating correlation in participatory value evaluation. *Environment and Planning B: Urban Analytics and City Science*, *0*(0), 23998083241264321. https://doi.org/10.1177/23998083241264321

Banna, J. C., Keim, N. L., & Townsend, M. S. (2023). Assessing face validity of a physical activity questionnaire for spanish-speaking women in california. *The Journal of Extension*, *49*(5), Article 7. https://doi.org/10.34068/joe.49.05.07

Bannigan, K., & Watson, R. (2009). Reliability and validity in a nutshell. *Journal of Clinical Nursing*, *18*(23), 3237–3243. https://doi.org/10.1111/j.1365-2702.2009.02939.x

Barbour, R. S. (2001). Checklists for improving rigour in qualitative research: A case of the tail wagging the dog? *BMJ*, *322*(7294), 1115–1117. https://doi.org/10.1136/bmj.322.7294.1115

Barnett, C., & Mahony, N. (2011). Segmenting publics.

Bickerstaff, K., & Walker, G. (2001). Participatory local governance and transport planning. *Environment and Planning A: Economy and Space*, *33*(3), 431–451. https://doi.org/10.1068/a33173

Bickerstaff, K., Tolley, R., & Walker, G. (2002). Transport planning and participation: The rhetoric and realities of public involvement. *Journal of Transport Geography*, *10*(1), 61–73. https://doi.org/https://doi.org/10.1016/S0966-6923(01)00027-8

Bouwmeester, M. (2021, August). *Effects of goal-dependent implementation choices on the achievement of goals in participatory value evaluation processes* [Master's thesis, Delft University of Technology] [to be defended publicly on Friday August 28, 2021 at 10:30 AM].

Boxebeld, S., Mouter, N., & van Exel, J. (2024). Participatory value evaluation (pve): A new preference-elicitation method for decision making in healthcare. *Appl Health Econ Health Policy*, *22*, 145–154. https://doi.org/https://doi.org/10.1007/s40258-023-00859-9

Buijs, A., & Boonstra, F. (2020). *Natuurbeleid betwist. visies op legitimiteit en het natuurbeleid*. KNNV Uitgeverij.

Burton, P. (2009). Conceptual, theoretical and practical issues in measuring the benefits of public participation. *Evaluation*, *15*(3), 263–284. https://doi.org/https://doi-org.tudelft.idm.oclc.org/10.1177/1356389009105881

Carlton, J., Peasgood, T., Mukuria, C., Connell, J., Brazier, J., Ludwig, K., Marten, O., Kreimeier, S., Engel, L., Belizán, M., Yang, Z., Monteiro, A., Kuharic, M., Luo, N., Mulhern, B., Greiner, W., Pickard, S., & Augustovski, F. (2022). Generation, selection, and face validation of items for

a new generic measure of quality of life: The eq-hwb. *Value in Health*, *25*(4), 512–524. https://doi.org/10.1016/j.jval.2021.12.007

Carreira, V., Machado, J. R., & Vasconcelos, L. (2016). Citizens' education level and public participation in environmental and spatial planning public policies: Case study in lisbon and surrounds counties. *International Journal of Political Science (IJPS)*, *2*(3), 25–34.

Cascetta, E., & Pagliara, F. (2013). Public engagement for planning and designing transportation system. *Procedia - Social and Behavioral Sciences*, *87*, 103–116.

Cascetta, E. (2009). Transportation systems analysis: Models and applications. https://api.semanticscholar.org/CorpusID:106443995

Coffé, H., & Bolzendahl, C. (2009). Same game, different rules? gender differences in political participation. *Social Science Quarterly*, *90*(5), 1205–1222. https://doi.org/10.1111/j.1540-6237.2009.00692.x

Columbia University. (2023). *Content analysis* [Accessed: [2 July 2024]]. https://www.publichealth.columbia.edu/research/population-health-methods/content-analysis

Connell, J., Carlton, J., Grundy, A., et al. (2018). The importance of content and face validity in instrument development: Lessons learnt from service users when developing the recovering quality of life measure (reqol). *Quality of Life Research*, *27*, 1893–1902. https://doi.org/10.1007/s11136-018-1847-y

Coulter, A., & Ellins, J. (2007). Effectiveness of strategies for informing, educating, and involving patients. *BMJ*, *335*(7609), 24–27. https://doi.org/10.1136/bmj.39246.581169.80

Dalawi, I., Isa, M., Chen, X. W., Azhar, Z., & Aimran, N. (2023). Development of the malay language - content validity face validity analysis of muaphq c-19. *BMC Public Health*, *23*, Article number. https://doi.org/10.1186/s12889-023-16044-5

Delaney, L. (2010). Descriptive statistics: Simply telling a story. *African Journal of Midwifery and Women's Health*, *4*(1). https://doi.org/10.12968/ajmw.2010.4.1.46313

Delgado, A., Lein Kjølberg, K., & Wickson, F. (2011). Public engagement coming of age: From theory to practice in sts encounters with nanotechnology. *Public Understanding of Science*, *20*(6), 826–845. https://doi.org/https://doi-org.tudelft.idm.oclc.org/10.1177/0963662510363054

de Ruijter, A. (2022). *[unpublished article]*.

de Ruijter, A. (2023). *[unpublished article]*.

Engel, L., Kosowicz, L., Bogatyreva, E., Batchelor, F., Devlin, N., Dow, B., Gilbert, A. S., Mulhern, B., Peasgood, T., & Viney, R. (2023). Face validity of four preference-weighted quality-of-life measures in residential aged care: A think-aloud study. *The Patient - Patient-Centered Outcomes Research*, *16*, 655–666. https://doi.org/10.1007/s40271-023-00647-6

Frogren, J., Schmidt, S. M., Kylén, M., Jonsson, O., Slaug, B., & Iwarsson, S. (2022). Awareness of and attitudes towards public involvement in research on ageing and health among older people in sweden. *PLoS ONE*, *17*(6), e0269993. https://doi.org/10.1371/journal.pone.0269993

Gaber, J., & Gaber, S. L. (2010). Using face validity to recognize empirical community observations [Challenges in Evaluation of Environmental Education Programs and Policies]. *Evaluation and Program Planning*, *33*(2), 138–146. https://doi.org/https://doi.org/10.1016/j.evalprogplan.2009.08.001

Ghauri, P. N., Gronhaug, K., & Strange, R. (2020). *Research methods in business studies*. Cambridge University Press.

Gigliotti, L., & Dietsch, A. (2014). Does age matter? the influence of age on response rates in a mixed-mode survey. *Human Dimensions of Wildlife*, *19*(3), 280–286. https://doi.org/10.1080/10871209.2014.895052

Glass, D., Kelsall, H., Slegers, C., Forbes, A., Loff, B., Zion, D., & Fritschi, L. (2015). A telephone survey of factors affecting willingness to participate in health research surveys. *BMC Public Health*, *15*, 1017. https://doi.org/10.1186/s12889-015-2350-9

Glenk, K., Meyerhoff, J., Colombo, S., & Faccioli, M. (2024). Enhancing the face validity of choice experiments: A simple diagnostic check. *Ecological Economics*, *221*, 108160. https://doi.org/10.1016/j.ecolecon.2024.108160

Golan, E. (2023, August). *Step into the driver's seat: A participatory value evaluation of the public transport policy preferences of the tel aviv metropolitan area & israeli face validity analysis* [Master's thesis, Delft University of Technology] [to be defended publicly on Thursday August 31, 2023 at 15:00 PM].

Hadi, N., Shafidan, N., Razali, S., Abdulateef, Q., & Hamid, M. (2023). The face and content validity of an instrument for measuring financial risk tolerance. *Journal of Computational Innovation and Analytics (JCIA)*, *2*, 57–88. https://doi.org/10.32890/jcia2023.2.1.4

Harris, J., Graue, M., & Dunning, T. (2018). Searching for the impact of participation in health and health research. *BioMed Research International*, *2018*.

Hermans, K., Spruytte, N., Cohen, J., Audenhove, C. V., & Declercq, A. (2016). Usefulness, feasibility and face validity of the interrai palliative care instrument according to care professionals in nursing homes: A qualitative study. *International Journal of Nursing Studies*, *62*, 90–99. https://doi.org/10.1016/j.ijnurstu.2016.07.014

Huijts, N. M. A., de Vries, G., & Molin, E. J. E. (2019). A positive shift in the public acceptability of a low-carbon energy project after implementation: The case of a hydrogen fuel station. *Sustainability*, *11*(8). https://doi.org/10.3390/su11082220

IBM. (2024). Crosstabulation Statistics in SPSS [Accessed: 2024-07-07].

Juschten, M., & Omann, I. (2023). Evaluating the relevance, credibility and legitimacy of a novel participatory online tool. *Environmental Science Policy*, *146*, 90–100. https://doi.org/https://doi.org/10.1016/j.envsci.2023.05.001

Knofczynski, G. T., & Mundfrom, D. (2007). Sample sizes when using multiple linear regression for prediction. *Educational and Psychological Measurement*, *68*(3), 431–442. https://doi.org/10.1177/0013164407310131

Kroesen, M. (2024). Lecture 6: Latent class cluster models [Accessed: August 7, 2024].

Liang, J., Bi, G., & Zhan, C. (2020). Multinomial and ordinal logistic regression analyses with multi-categorical variables using r. *Annals of Translational Medicine*, *8*(16), 982. https://doi.org/10.21037/atm-2020-57

Mason, J., et al. (2020). Establishing face and content validity of a survey to assess users' perceptions of automated vehicles. *Journal of Ergonomics*, *8*(3), 10–15. https://doi.org/10.1080/00140139.2020.1800001

Meyer, M. D. (2016, August). *Transportation planning handbook* (4th). ITE (Institute of Transportation Engineers), John Wiley & Sons, Inc.

Ministerie van Infrastructuur en Waterstaat. (2024). MIRT-verkenning Oude Lijn [Accessed: 2024-07-01]. https://www.mirttrajecten.nl/organisatie/oude-lijn

Mohd Pakri, M. A., Arif Shah, J., & Mohamed Haris, N. B. (2024). The perspective of vegetable farmers in readiness to face the post covid-19 pandemic: A cross-tabulation analysis. *Adv Agri Food Res J*, *5*(1), a0000502. https://doi.org/10.36877/aafrj.a0000502

Morris, A., Onwuegbuzie, A. J., & Gerber, H. R. (2018). Using expert interviews within modes in online and offline spaces to extend comprehensive literature review processes. *The Qualitative Report*, *23*(8), 1777–1798. https://doi.org/10.46743/2160-3715/2018.3039

Morton, S., Baccini, A., & Morandi, A. (2023). Content and face validity of the evaluation tool of health information for consumers (ethic). *International Journal of Environmental Research and Public Health*, *20*(6), 4872. https://doi.org/10.3390/ijerph20064872

Mouter, N., Hernandez, J., & Itten, A. (2021). Public participation in crisis policymaking. how 30,000 dutch citizens advised their government on relaxing covid-19 lockdown measures. *PLoS ONE*, *16*(5), e0250614. https://doi.org/10.1371/journal.pone.0250614

Nevo, B. (1985). Face validity revisited. *Journal of Educational Measurement*, *22*(4), 287–293. Retrieved March 1, 2024, from http://www.jstor.org/stable/1434704

Oladeji, S. O., et al. (2022). Community participation in conservation and management of cultural heritage resources in yoruba ethnic group of south western nigeria. *Journal of Heritage Management*, *10*, X–X.

Oluwatayo, J. A. (2012). Validity and reliability issues in educational research. *Journal of Educational and Social Research*, *2*(2), 391–391. https://doi.org/10.5901/jesr.2012.v2n2.391

Onyango, L. A., Mutui, F. N., & Wabwire, E. (2019). An analysis of the factors affecting public participation in environmental impact assessment: Case study of selected projects in nairobi city county, kenya. *European Scientific Journal*, *15*(9), 284–301. https://doi.org/10.19044/esj.2019.v15n9p284

Öztürk, S. (2023). Face validation in large-scale simulation projects. *Mugla Journal of Science and Technology*, *9*(1), 24–33. https://doi.org/10.22531/muglajsci.1179857

Patel, N., & Desai, S. (2020). Abc of face validity for questionnaire. *International Journal of Pharmaceutical Sciences Review and Research*, *65*, 164–168. https://doi.org/10.47583/ijpsrr.2020.v65i01.025

Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd). Sage Publications.

Populytics. (2021). *Rapport-raadpleging optimalisatie stelsel medische rijgeschiktheid* (tech. rep.) (Accessed: 2024-07-24). Populytics. https://populytics.nl/

Quick, K. (2014, August). Public participation in transportation planning.

Reffien, M. A. M., Ismail, A., Sulong, S., Bahanuddin, T. P. Z. T., Salleh, N. M., & Nasir, N. H. (2022). Development, validity and reliability of an instrument to evaluate implementation fidelity of the family doctor concept programme in public primary care clinics in malaysia. *Malaysian Family Physician*, *17*(1), 20–28. https://doi.org/10.51866/oa1197

Rijkwaterstaat. (2022). *Get started with broad prosperity. an exploration of how rijkswaterstaat can manage broad prosperity with existing instruments* [Performed by Waai].

Rosener, J. B. (1978). Citizen participation: Can we measure its effectiveness? *Public Administration Review*, *38*(5), 457. https://doi.org/10.2307/975505

Rowe, G., & Frewer, L. J. (2004). Evaluating public-participation exercises: A research agenda. *Science, Technology, & Human Values*, *29*(4), 512–556. https://doi.org/10.1177/0162243903259197

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, *104*, 333–339.

Spruit, S., & Mouter, N. (2020). 1376 inwoners van súdwest-fryslân over het toekomstige energiebeleid van hun gemeente: De uitkomsten van een raadpleging.

Taherdoost, H. (2016). Validity and reliability of the research instrument; how to test the validation of a questionnaire/survey in a research (august 10, 2016).

Tuit, C. (2022). *Face validity of the participatory value evaluation method: A case study of the schiphol environmental council* [Master's thesis, Delft University of Technology] [Unpublished master's thesis].

Van Wee, B., Annema, D., & Banister, D. (2013). *The transport system and transport policy, an introduction*. Edward Elgar.

van Infrastructuur en Waterstaat, M. (2020). Hoofdlijnennotitie mobiliteitsvisie 2050 [Accessed: 2024-07-01]. https://www.rijksoverheid.nl/documenten/rapporten/2020/09/01/bijlage-hoofdlijnennotitie-mobiliteitsvisie-2050

Vermunt, J. K., & Magidson, J. (2002). Latent class cluster analysis. In J. A. Hagenaars & A. L. McCutcheon (Eds.), *Applied latent class analysis* (pp. 89–106). Cambridge University Press.

Wester, M., & Morn, C. (2013). Gender and public participation: A study of the swedish public debate on the nord stream gas pipeline from a gender perspective. *Journal of Environmental Assessment Policy and Management*, *15*(03), 1350016. https://doi.org/10.1142/S1464333213500166

Zheng, Y., & Schachter, H. L. (2016). Explaining citizens' e-participation use: The role of perceived advantages. *Public Organization Review*, *17*(3), 409–428. https://doi.org/10.1007/s11115-016-0346-2

A Scientific Paper

Face Validity in Participatory Value Evaluation

Exploring Segment-specific Perceptions and their Influence on Transport Decision-Making

A.R. Ramadani
Delft University of Technology

***Abstract*—Transportation planning is increasingly incorporating public participation, with Participatory Value Evaluation (PVE) emerging as a key method. However, the validity of PVE, particularly its face validity across different demographic groups, remains under-explored. This study addresses this gap by analyzing face validity evaluations from three recent PVE consultations in the Netherlands. Statistical analyses, content analysis, and expert discussions were employed to examine how factors such as age, gender, and education influence these evaluations. The findings indicate that while PVE is generally perceived positively, specific groups—particularly older adults and those with lower education levels—consistently evaluate its face validity negatively. To enhance the legitimacy of PVE in transportation planning, tailored strategies are recommended to address these persistent negative evaluations.**

***Index Terms*—Participatory Value Evaluation (PVE), Face Validity, Public Participation, Transport Appraisal, Transportation Planning.**

I. Introduction

Transportation serves as a fundamental pillar of modern society, facilitating the movement of individuals, goods, and services for purposes such as employment, commerce, and social interactions [35]. However, alongside these benefits, transportation activities also introduce challenges, including environmental pollution and safety concerns. Effective transportation planning is crucial to maximizing the benefits of transport while mitigating its negative impacts.

Transportation planning encompasses a wide range of activities, from the development of transport plans, policies, and projects to their implementation. [36] outlines these activities, spanning short-term operational management, tactical planning, feasibility studies, and strategic long-term planning, often involving collaboration with various stakeholders.

Historically, transportation projects have followed a systematic process involving goal setting, problem identification, the generation of alternatives, and evaluation. Tools such as Social Cost-Benefit Analysis (SCBA) have been employed to comprehensively assess the impacts of policy interventions on all stakeholders [35]. In recent decades, there has been a significant shift toward more inclusive and deliberative decision-making processes. Traditional approaches have been criticized for neglecting the perspectives of residents, thereby hindering public participation [38]. Emerging studies advocate for participatory approaches, highlighting their potential to enhance social cohesion, fairness, service quality, and societal learning [39]. These approaches are seen as leading to more legitimate policy decisions [37], a transformation reflected not only in transportation planning but also in other domains of public decision-making. Consequently, legislative frameworks, particularly in democratic nations, have evolved to support principles of democracy and inclusiveness in transportation planning [40].

In transportation planning, citizen involvement is increasingly realized through digital participatory approaches, facilitated by advancements in technology. These methods enable effective communication, empower citizens, and reduce costs in policy-making processes [41]. Among these methods, Participatory Value Evaluation (PVE) has emerged as a promising approach.

PVE is an online experiment that allows citizens to engage in the decision-making process from the perspective of policymakers. Participants are provided with information about various policy options, including their characteristics and potential effects, and are required to consider factors such as government budget constraints and sustainability objectives when making choices. These preferences reflect both personal utility maximization and societal values [42]. In practice, PVE is typically initiated before final policy decisions are made, with participants able to contribute further input in subsequent phases. The objective is to generate recommendations for policymakers based on the insights gained through the PVE process [43].

The effectiveness of public participation hinges on a thorough assessment of its validity [44], [45]. Validity, which measures the extent to which an instrument accurately captures the intended construct, is multifaceted and context-dependent [1], [46], [47]. Despite its importance, validity is often overlooked in commonly used instruments due to various constraints [32], [48]. However, neglecting validity checks, particularly in instruments like PVE consultations, can lead to misleading results and misguided decisions [49].

Face validity, one of the quickest methods to assess an instrument, evaluates whether, on the surface, the instrument appears logical and connected to the intended concept [2]. Although sometimes criticized as 'trivial' or 'cosmetic,' face validity remains an essential component of citizen participation processes, as it ensures the perceived authenticity and utility of the method, making it a legitimate tool for policymakers [48], [50].

Despite the significance of face validity, its application in PVE remains under-explored. Understanding why participants evaluate the face validity of PVE is critical, particularly in terms of inclusiveness and sustainability in transportation

planning. This requires segmentation of respondents based on demographics, behaviors, or attitudes to tailor engagement strategies and improve the face validity of PVE. By analyzing participation results at a granular level, organizations can identify trends, preferences, and concerns specific to each segment, thereby enhancing decision-making [51]. Identifying groups that consistently evaluate PVE negatively allows for the development of tailored strategies in future PVE designs to increase its face validity and effectiveness as a public participation tool.

This paper contributes to the scientific discourse by investigating the validity of the PVE method, an area that has received limited attention. Specifically, it explores the face validity of PVE, shedding light on citizens' expectations regarding public participation within the PVE framework. This insight enables future PVE initiatives to better accommodate citizens, allowing for more informed advice to policymakers and improving the overall effectiveness of their participation. The primary objective of this research is to answer the research question *"To what extent are there distinct segments in the population who (in)consistently evaluate the face validity of Participatory Value Evaluation (PVE) in transportation planning?"* and to propose recommendations for accommodating these groups in future PVE initiatives. By validating the PVE method, this research enhances its credibility and utility as a decision-making tool, ultimately enriching the practice of transportation planning.

The subsequent sections of this paper are organized as follows: Section II details the cases and methods used in this research, while Section III presents the results of the literature review. Section IV discusses the hypotheses, while Section V discusses the findings of both quantitative and qualitative analyses conducted in this study. Following that, Section VI provides discussion and recommendations, and finally, Section VII concludes the paper.

II. Methodology

This study utilized a multi-method approach, which included doing a literature review, statistical analyses, content analyses and group discussion with experts, in order to thoroughly evaluate the face validity of Participatory Value Evaluation (PVE). These methodologies were chosen to address both the quantitative and qualitative parts of the research. The conceptual framework of research methodology of this study is presented in Table 1.

A. Literature Review

The literature review focused on identifying existing studies on public participation, PVE, and research instrument validity. Keywords such as 'citizen participation,' 'PVE,' instrument validity', and 'face validity' were used in databases like Scopus and Google Scholar to gather relevant studies, which informed the design of the study and identification of research gaps. Additionally, the master theses of [1] and [3] were utilized for backward snowballing to provide an overview of the current condition of PVE and face validity. Forward snowballing were also undertaken from [2] and [46] to determine the current research on validity of research instrument. Articles were chosen based on their relevance to public participation in decision-making, particularly the PVE approach in transportation, and the validity of preference elicitation experiments.

B. Case Studies and Data Collection

The data for this study were obtained from three Participatory Value Evaluation (PVE) consultations that were recently carried out in the Netherlands. The consultations were chosen based on their extensive and recent datasets, which specifically address key transportation and accessibility topics that are in line with the Netherlands government's objectives. The datasets are specified in Table 1, and were selected based on their relevance to the research objectives and the consistency of their outcomes, as evidenced in the descriptive results (refer to Section V-A2).

Table 1: PVE Consultations Utilized in this Study

No	Year	Consultation	Area	Project in Consultation
1	2023	Lelylijn	North Holland	Railway line that links the northern part of the Netherlands with the rest of the country
2	2024	Oude Lijn	South Holland	Regional development along the existing railway "Old" Line
3	2024	Mobility Vision	Netherlands	Government's long-term plan for mobility from now until 2050

The three PVE consultations in Table 1 were intentionally designed with slight variations to cater to the distinct objectives of each consultation. Each consultation included several consistent sections: policy selection and rationale, sociodemographic data collection, participant reviews of the consultation, and an additional section tailored to the specific objectives of each consultation. For the purposes of this study, only the sociodemographic data and participant reviews were utilized, focusing on their evaluations of the face validity of PVE. The face validity statements that were used for the evaluation in the study are outlined in Table 2.

C. Data Cleaning and Processing

The datasets acquired from the three PVE consultations contained extensive information, much of which was not relevant for this study. Therefore, a data cleaning process was undertaken to retain only the necessary variables: respondents' sociodemographic information, their assessments of face validity, and their overall ratings of the PVE consultation. For the Lelylijn dataset, additional responses related to what participants did not like about the consultation were retained for subsequent content analysis (see section V-B). Responses that included "I don't know" or "I prefer not to say" options were excluded to ensure consistency across the datasets. Additionally, only complete responses were included in the analysis—if a respondent failed to answer any of the seven face validity questions, their entire response was removed. As a result, the total number of responses and the descriptive

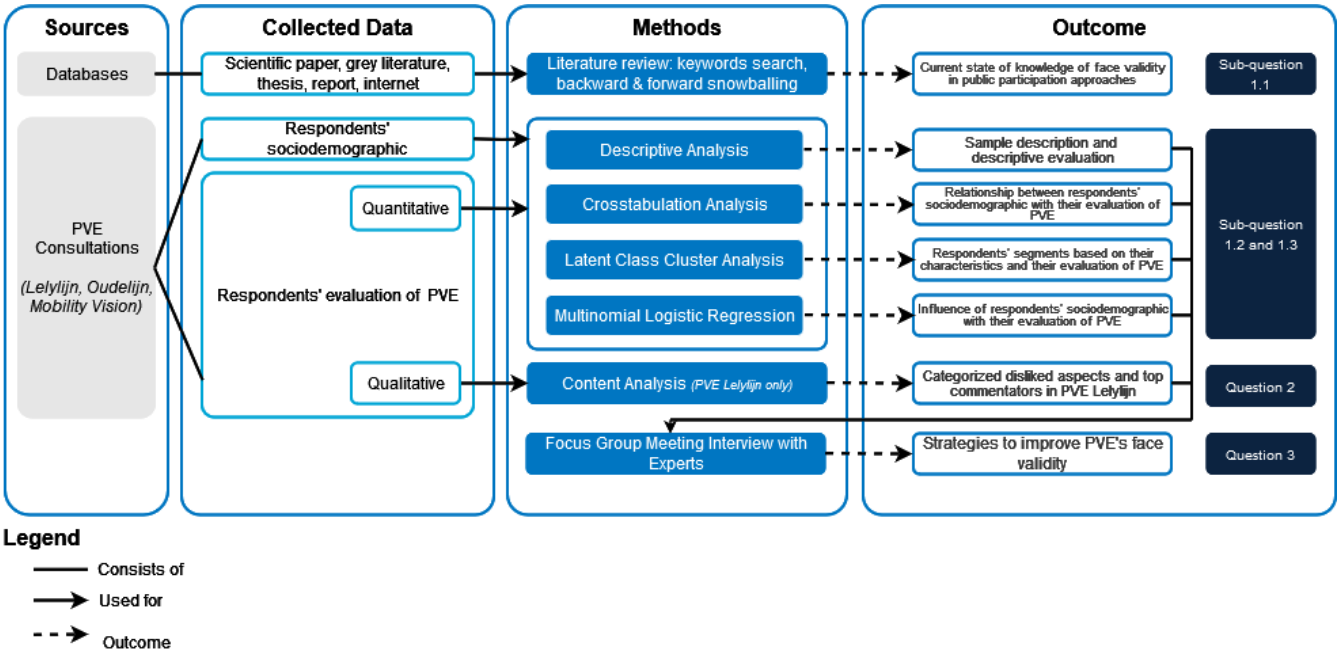


Fig. 1: Conceptual Framework of Research Methodology

Table 2: Face Validity Statements in PVE Consultations

No	Dimension	Face Validity Statement	Lelylijn	Oude Lijn	Mob. Vision
1	Transparency	The consultation steered my choices in a certain direction	✓	✗	✗
2	Transparency	I trust this is an honest investigation	✗	✓	✓
3	Feasibility	I found the consultation difficult to understand	✓	✓	✓
4	Relevance	I thought it was an important topic to give my opinion	✓	✓	✓
5	Relevance	By participating in this consultation, I have learned about the choices that the government has to make on this subject	✓	✓	✓
6	Acceptance	This method should be used more often to involve residents in public policy	✓	✓	✓
7	Acceptance	If many people participate in this consultation, then the final decisions on this subject are better accepted for me	✓	✓	✓
8	Acceptance	If the government often allows residents to think along about these types of choices more often, then I get more confidence in the decisions of the government	✓	✓	✓

statistics presented in this study may differ from those in the original PVE consultation reports

Following data cleaning, descriptive analysis revealed demographic discrepancies between the respondent profiles and the general Dutch population, particularly in terms of gender, age, and education. To address this, factor weighting was applied to adjust for these discrepancies in subsequent analyses. The factor weighting process involved calculating Chi-square statistics to assess the alignment between the sample proportions and the population proportions. Iterative adjustments were made until the sample proportions were statistically comparable to those of the Netherlands population.

D. Statistical Analysis

Following the data cleaning and weighting process, the dataset was prepared for statistical analysis using SPSS and LatentGold 6.0 for Latent Class Cluster Analysis. The statistical methods employed are detailed below:

1) *Descriptive Analysis*
Descriptive analysis was utilized to characterize the population sample and provide an overview of the collected data, including the distribution of sociodemographic characteristics, face validity responses, and overall consultation ratings [6]. This analysis helped establish a baseline understanding of the dataset before proceeding with more complex analyses..
2) *Crosstabulation Analysis*
Crosstabulation analysis was conducted to explore the relationships between respondents' sociodemographic characteristics and their evaluations of PVE's face validity, as well as their overall consultation ratings. In this analysis, sociodemographic characteristics (e.g., age, gender, education) were placed in the rows, while face validity ratings and overall consultation evaluations were placed in the columns. This approach allowed for the examination of associations between the different categories within these variables.

The Gamma coefficient was employed to measure the strength and direction of these associations, with values ranging from -1 to 1. A positive Gamma value indicates a direct relationship, while a significance (Sig.) value below 0.05 was used to determine statistical significance, indicating that the observed relationships are unlikely to be due to chance.

3) *Latent Class Cluster Analysis (LCCA)*
Latent Class Cluster Analysis (LCCA) was performed to identify distinct subgroups within each consultation that share similar respondents' characteristics and PVE evaluation patterns. Positive evaluations were defined by agreement with positive face validity statements and high consultation ratings, while negative evaluations were characterized by disagreement with positive statements or low ratings.
The analysis used categorical face validity responses ("agree," "neutral," "disagree") as indicators, alongside overall consultation ratings. Sociodemographic characteristics were included as covariates. Models with 1 to 10 clusters were tested, with the optimal model selected based on several criteria: Log-Likelihood (LL) for model fit, Information Criteria (BIC, AIC, AIC3) for model parsimony, Maximum Bivariate Residual (Max. BVR) for pairwise association discrepancies, classification error rates for accuracy, and Entropy R^2 for cluster distinction clarity. The interpretability and meaningfulness of the clusters were also key considerations.
4) *Multinomial Logistic Regression*
Initially, Ordinal Logistic Regression was considered for examining the influence of respondents' characteristics on face validity evaluations across ordered categories. However, parallel line tests indicated a violation of the proportional odds assumption, necessitating an alternative approach [7].
Consequently, Multinomial Logistic Regression was employed to model the relationship between respondents' characteristics and their face validity ratings, categorized as "agree," "neutral," and "disagree." This method accommodates non-linear relationships without requiring the proportional odds assumption. Independent variables included demographic factors (age, gender, education) and case-specific characteristics (e.g., proximity to the study site, prior PVE experience).

E. Content Analysis

Content analysis was employed to examine the textual data from the Lelylijn dataset, utilizing both qualitative and quantitative methods [8]. This approach was crucial for gaining a deeper understanding of the reasoning behind participants' assessments of the PVE consultation, particularly their critiques. The analysis process involved categorizing respondents' comments into specific themes, which were then aligned with relevant dimensions of face validity, such as transparency, feasibility, and acceptance. This thematic categorization allowed for a nuanced interpretation of the data, providing insights into the specific aspects of the consultation that participants found unfavourable.

While content analysis is recognized as a time-intensive method that carries a risk of potential bias, it offers valuable insights by combining qualitative depth with quantitative rigor [9]. Given the complexities and resource demands of content analysis, this method was selectively applied to the Lelylijn negative feedback dataset, which was both the first consultation conducted and the earliest available, providing timely access to the richest and most detailed qualitative feedback.

F. Focus Group Meeting Interview with Experts

Expert interviews were conducted to gain contemporary insights and deepen the understanding of the challenges and opportunities associated with Participatory Value Evaluation (PVE) consultations [10]. The primary objective was to explore practical solutions for improving public perceptions of PVE, particularly in transportation, where PVE is a relatively new approach.

The focus group consisted of nine experts selected based on their extensive experience and significant contributions to the field of PVE consultation design and research. The group included university researchers, independent scholars, and professionals who have developed PVE methodologies, published extensively on PVE and transport appraisal, and worked closely with government agencies on policy-making. Selection criteria included their involvement in significant PVE projects, the relevance of their work to transportation planning, and their recognition within the academic and professional communities.

The focus group meeting began with a presentation of the quantitative and qualitative findings from the previous stages of the study. This was followed by a Q&A session where experts could seek clarification on the findings. The central question posed to the group was: *"How can we address the negative evaluations given by certain groups?"* This open-ended question facilitated a broad and in-depth discussion, allowing participants to share their opinions and propose potential solutions. The discussion was moderated to ensure that all experts had the opportunity to contribute, and key points were recorded for subsequent analysis. The insights gathered were then synthesized and integrated into the study's recommendations, ensuring that the final strategies were informed by both empirical data and expert opinion.

III. Literature Review

Validity is a fundamental aspect in the construction and evaluation of measurement instruments, ensuring that they accurately capture the specific subject under study [11]. Various forms of validity are recognized in the literature, including face validity, content validity, construct validity, and criterion validity [2]. Among these, face validity is particularly crucial in public participation methods, such as PVE, where the perceived credibility of the process significantly influences its acceptance and effectiveness [12].

Face validity refers to the extent to which a measurement tool appears to measure what it claims to, based on its relevance, clarity, and acceptability to the target audience [1]. It is critical because when participants perceive an instrument as having high face validity, they are more likely to engage meaningfully and provide accurate responses. This enhances the credibility and acceptance of the instrument, making it an essential component in fields that rely on public input.

Public participation initiatives offer significant benefits, such as enhancing participants' dignity, self-esteem, and social engagement, by involving them in decision-making processes [13]. However, these benefits are contingent upon the perceived validity of the participation processes and instruments used. Ensuring face validity is vital for building trust, fostering engagement, and improving the quality of public participation outcomes. Without perceived validity, even the most well-intention participation efforts may fail to achieve their objectives.

In transportation sector, PVE is a relatively new method for gathering public preferences. [1] introduced a framework for assessing the face validity of PVE, which includes statements that evaluate the method's clarity, relevance, and comprehensibility. A study by [3], has applied this framework to assess face validity in PVE about public transport policies. Further, [14] developed a specific face validity assessment tool for PVE, identifying nine key dimensions that are critical to evaluating the method's effectiveness. These dimensions serve as a foundation for understanding how different aspects of PVE are perceived by participants, providing valuable insights for practitioners and policymakers.

Despite the growing body of literature on face validity and its importance in public participation, there is a notable gap in understanding how different demographic groups (in)consistently evaluate the face validity of participation approaches like PVE. Previous studies have largely focused on the general population or specific policy contexts, without adequately exploring demographic variations in perceptions of validity. This study seeks to address this gap by employing a mixed-methods approach to analyze face validity evaluations across different demographic groups, using data from recent PVE consultations in the Netherlands.

IV. Hypotheses

Based on existing literature, this study proposes the following hypotheses regarding the evaluation of Participatory Value Evaluation (PVE) consultations in the context of transportation planning:

1) Older individuals tend to evaluate PVE more positively than younger individuals. Older adults are more likely to engage positively with online participatory methods like PVE due to a higher sense of political efficacy and a greater willingness to participate in online activities compared to younger individuals [52], [53].
2) Women generally have a more positive attitude towards PVE compared to men. Women are often more supportive of participatory processes and are more likely to engage in informal, private political activities, which aligns with the nature of PVE [54], [55].
3) Highly educated respondents view PVE positively but with a more critical perspective. Individuals with higher education levels are more engaged in public policy discussions and possess the skills to critically assess the effectiveness of participatory methods like PVE [57], [58].
4) Frequent users of the train or the specific project tend to evaluate PVE more positively. Frequent users of transportation systems are likely to have a more favorable view of PVE, as familiarity with the subject matter enhances their engagement and positive perception [59].
5) Individuals living closer to the project are more likely to perceive PVE positively. Proximity to the project increases the perceived benefits and value of participation, leading to a more positive evaluation of PVE [60], [61].

V. Results

A. Evaluation of Face Validity in PVE Consultations

1) Sample Characteristics:

This subsection presents the sociodemographic characteristics of participants across the three PVE consultations. Table 3 includes both the attributes used for factor weighting (age, gender, and education) and additional attributes relevant to the analysis (e.g., occupation, proximity to the project, frequency of train use). The factor weighting process, as detailed in Section II-C, was applied to all attributes, ensuring that the dataset is representative of the Dutch population in terms of key sociodemographic factors.

Table 3 outlines the sample distribution across various attributes, showing how closely the weighted sample aligns with the Dutch population for the characteristics used in factor weighting (gender, age, and education). Other characteristics, such as occupation, proximity to the project, and frequency of train use, are also included in the analysis but were not part of the factor weighting process due to the lack of corresponding population data. Despite this, factor weighting was applied across the entire dataset to ensure that all attributes are appropriately adjusted for representativeness. For further details on the methodology and characteristics of the sample, readers are referred to [62], which provides an in-depth analysis of these attributes across all three datasets.

2) Descriptive Results:

Descriptive statistics were employed to assess respondents' ratings of both the overall and face validity of the PVE consultations. Each consultation featured seven face validity statements, which participants rated using a Likert scale (ranging from "totally disagree" to "totally agree"). While the phrasing of these statements varied slightly across consultations, their core content remained consistent (see Table 2).

The results of the descriptive analyses for each consultation indicated a generally positive appraisal of face validity in all three PVE consultations. According to the analysis of negatively formulated statements shown in Fig. 2a and Fig. 2b, about 46.8% of respondents did not think that the

Table 3: Sample (proportion%) After Factor Weighting vs. Population Distribution in the Netherlands. Empty cells Means Specific Attribute or Level is not Available in the Study.

Respondents' Attribute	Lelylijn	Oude Lijn	Mobility Vision	Population Nether-lands
Gender				
Man	2,332 (51%)	2,049 (50%)	2,161 (51%)	51%
Woman	2,240 (49%)	2,021 (50%)	2,097 (49%)	49%
Age				
18 to 34 years	1,215 (27%)	1,118 (27%)	1,134 (27%)	27%
35 to 64 years	2,220 (49%)	1,964 (48%)	2,041 (48%)	48%
65 years or older	1,137 (25%)	987 (25%)	1,083 (25%)	25%
Education Level				
Low	1,245 (27%)	1,115 (27%)	1,190 (28%)	29%
Middle	1,684 (37%)	1,508 (37%)	1,565 (37%)	37%
High	1,643 (36%)	1,447 (36%)	1,502 (35%)	35%
Occupation				
Not working	1,773 (39%)	1,638 (40%)	1,671 (39%)	
Part-time	1,022 (22%)	868 (21%)		
Full-time	1,778 (39%)	1,564 (38%)	2,587 (61%)	
Residency				
Tenant	975 (21%)	446 (11%)		
Lodger	395 (9%)	1,068 (26%)		
Homeowner	3,202 (70%)	2,555 (63%)		
Proximity				
Not proximate	2,312 (51%)	1,915 (47%)		
Proximate	2,261 (49%)	2,154 (53%)		
Freq. Train Use				
Rarely	2,869 (63%)	2,240 (55%)		
Occasionally	925 (20%)			
(Almost) daily	778 (17%)	1,829 (45%)		
Freq. Line Use				
Rarely	2,650 (58%)	2,748 (68%)		
Occasionally	1,260 (28%)			
(Almost) daily	663 (15%)	1,322 (32%)		
Area of Living				
Urban		2,810 (69%)	2,610 (61%)	
Rural		1,259 (31%)	1,647 (39%)	

consultation influenced their choices. Meanwhile, over 70% of respondents find this consultation to be understandable, in both the Oude Lijn and Mobility Vision consultations. However, a notable proportion of respondents remain impartial (37.4%) and consider the Lelylijn consultation as steering their decisions. Additionally, approximately 9.7% of respondents in Oude Lijn and Mobility Vision think that the consultation is difficult to understand, highlighting an area that warrants further attention in future PVE designs.

The statement is *"I trust this is a fair investigation,"* was included in all three consultations, and over 80% of the respondents expressed confidence in the integrity of the consultation (see Fig. 2c). Notably, less than 3% of respondents in Oude Lijn and Mobility Vision believe that this consultation lacks honesty, while solely 5.8% of respondents in the Lelylijn consultation expressed doubts about its truthfulness.

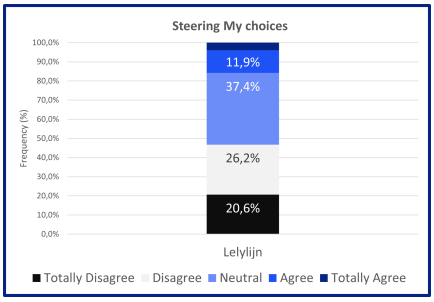
In the relevance dimension, as shown in Fig. 2d, there are two statements that are present in all three PVE consultations. According to the statement, 93% of the respondents of Lelylijn considered the consultation important, likely because the project has not yet been implemented, making it crucial for them to voice their opinions. The finding is noteworthy due to the fact that a majority of the participants (54.3%) in the Lelylijn consultation live far from the project. The Oude Lijn consultation, although recognized to be important by a large proportion of individuals, had the least consensus regarding its importance (72.8%) compared to the other two consultations. One possible reason for this is because although the consultation asked for suggestions for improvements and changes, the Oude Lijn already exists and is now able to provide "sufficient" services to the respondents.

Regarding the statement *"By participating in this consultation, I have learned about the choices that the government has to make on this subject"*, 70.3% of respondents in both the Lelylijn and Oude Lijn believe that they learned about the government's decisions through the consultation. In the Mobility Vision consultation, a lower portion of respondents of 61.4%, (totally) agreed with the statement. However, disagreed with this statement, and at least 20.0% remained neutral across all three consultations. This suggests that a significant portion of participants did not feel adequately informed about the government's decisions, indicating an opportunity for future PVE designs to enhance the informational content provided to respondents.

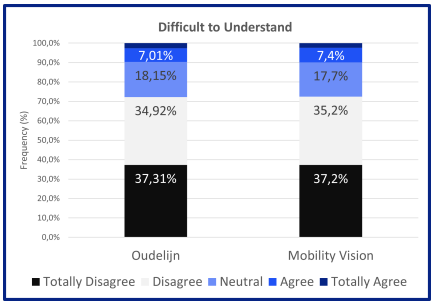
In the acceptance dimension, depicted in Fig. 2f, at least 79% of the participants in three consultations (totally) agreed with the initial statement *"This method should be used more often to involve residents in government policies"*. A mere 3.4% of the respondents expressed their disagreement, indicating that respondents broad support for the use of PVE in public policy engagement. The statement *"If many people participate in this consultation, the final decisions on this topic are more acceptable to me,"* similarly elicits a significant level of agreement, with at least 61% of the respondents throughout the three consultations expressing agreement. Nevertheless, a minimum of 6.6% disagreed, and 22.4% remain indifferent, suggesting that while the majority of respondents are inclined to accept the final decisions on the subject, there is still a subset of individuals who are unlikely to increase their acceptance on the final decisions, regardless of the high number of participants in the consultation.

Finally, the statement *"If the government involves residents more often in thinking about these kinds of choices, I will have more confidence in the government's decisions"* exhibits similar patterns to the preceding statement regarding public receptiveness towards final decisions. Over 65.1% of the participants in all three consultations experience that participating in PVE increases their trust in the government. However, at least 6.7% of the participants express skepticism, expressed skepticism, indicating that while PVE generally fosters trust, there is still a segment of the population that remains unconvinced.

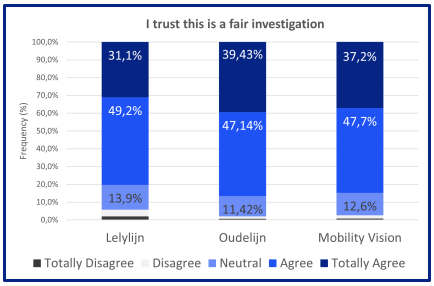
Respondents were also asked to rate the overall consultation. The average scores were 7.42 for Lelylijn, 7.72 for Oude Lijn, and 7.55 for Mobility Vision. As depicted in Fig. 3, the distribution of ratings across the consultations is relatively



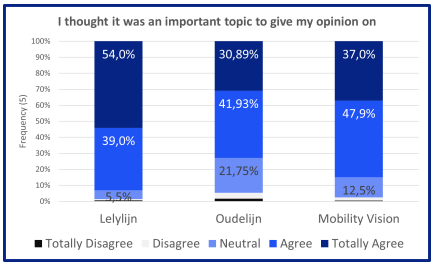
(a) Consultation is Steering



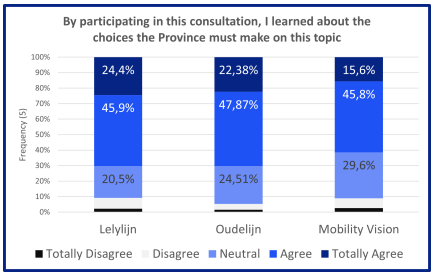
(b) Difficult to Understand



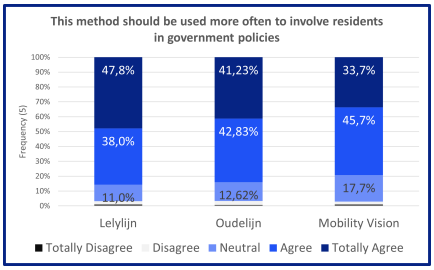
(c) Trust Investigation is Fair



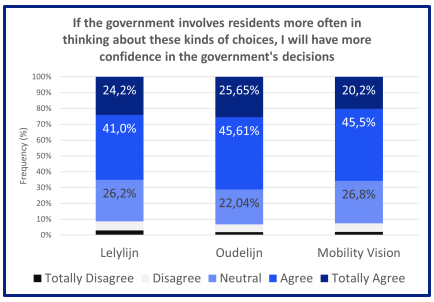
(d) Important Topic



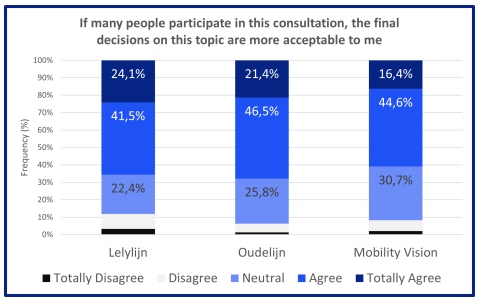
(e) Learned Government's Choices



(f) This Method Should be Used More



(g) Increase Acceptance Decisions



(h) Increase Trusts in Government

Fig. 2: Comparisons of Face Validity Evaluation across PVE consultations

similar, though the Oude Lijn consultation received slightly higher average ratings. Notably, Lelylijn had the lowest average rating, with a few respondents scoring it as low as 1 or 3, although these cases were not substantial.

3) Crosstabulation Analysis:

Crosstabulation analysis was conducted to explore the relationships between respondents' characteristics and their evaluations of face validity across the PVE consultations. Significant associations were found between various demographic factors and how respondents rated different aspects of the consultations, though these associations were not always consistent across the three consultations.

Table 4 summarizes the relationships between respondents' attributes and their evaluations of face validity statements. A positive sign (+) indicates a positive correlation between the characteristic (e.g., age, gender) and agreement with the face validity statement. A negative sign (-) indicates a negative correlation, while "NS" denotes a lack of significant relation-

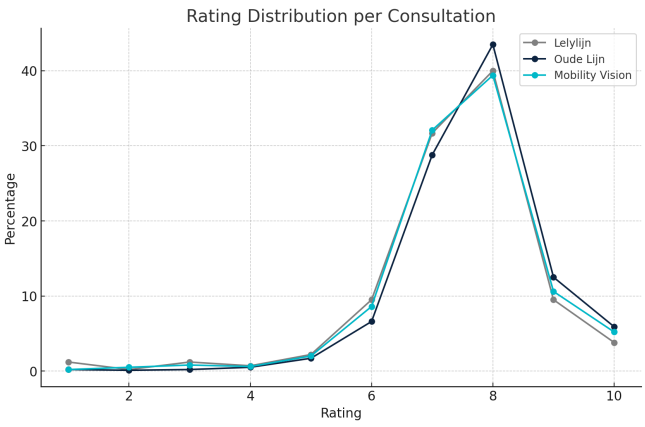


Fig. 3: Consultation Ratings Distribution Across 3 PVE Consultations

ship. ”NC” means the association was not checked for that consultation. For instance, in the Lelylijn consultation, the table shows a positive correlation (+) between proximity to the project and the perception that the consultation is steering choices, suggesting that respondents living closer to the project are more likely to feel influenced by the consultation process. The detailed results of the crosstabulation analysis for each consultation can be found in [62]'s thesis. The key comparisons among the three PVE consultations are summarized as follows:

Age consistently played a significant role in shaping perceptions. Older respondents were more likely to view the consultations as steering their decisions and found the process less difficult to understand. However, they were also less likely to perceive the consultations as fair or feel that they had learned about the government's choices. Interestingly, age did not consistently influence views on the importance of the topic or the frequency with which the method should be used.

Gender differences also influenced perceptions, although the patterns were less consistent. Men were generally less likely to feel that they gained insights into government decisions and were more likely to perceive the process as steering. However, the relationship between gender and perceptions of the consultation's importance and the desire to use the method more frequently varied depending on the specific consultation.

The level of education is another factor that significantly influences views on PVE consultations. Individuals with higher levels of education are generally less likely to experience PVE to increase their trusts in the governments. Conversely, more educated respondents found the consultation process easier and regarded the topics discussed as more important. However, inconsistencies emerged regarding how education associated with perceptions of fairness, knowledge gained, and the desirability of using the method more frequently.

Occupation showed mixed results in its influence on perceptions. Full-time workers were more likely to think that the consultation process was not overly steering. However, there was no consistent pattern in how occupation affected overall assessments of face validity, suggesting that the impact of occupation may depend on specific contextual factors.

Proximity to the project significantly influenced perceptions. Respondents living closer to the project were more likely to view the consultation as steering their decisions, yet they also tended to perceive the consultation as fair and the topic as important. These respondents were also more likely to support the increased use of this consultation method, although the consistency of these findings varied.

Frequent train users consistently viewed the consultations more positively. They were more inclined to find the consultations fair and important, felt better informed about government decisions, and were more supportive of increasing the use of the consultation method. These respondents were also less likely to perceive the process as steering and more likely to trust the government's decisions, although not all evidence supported these trends.

The frequency with which respondents planned to use the specific train line under discussion was positively associated with their perceptions of the consultations. Those who intended to use the line more frequently were more likely to find the consultations fair and important and were supportive of increasing the use of this method. Additionally, these respondents tended to feel more confident about the consultation process, believed they had gained knowledge, and were more accepting of government decisions, though evidence for these associations was not entirely consistent across consultations.

4) Latent Class Cluster Analysis (LCCA):

The Latent Class Cluster Analysis (LCCA) conducted for each PVE consultation revealed consistent results, identifying four distinct clusters in each case, as summarized in Table 5. Although the demographic characteristics of respondents in different clusters varied across consultations, the evaluations of face validity within each cluster showed notable similarities. It is important to note that although the group with greatesT probable outcomes per cluster in Table 5 may be comparable (e.g., Age), the actual percentages differ and can be examined in more detail in [62].

The largest cluster in each consultation generally gave the most positive assessments of the PVE consultations. This group primarily consisted of middle-aged individuals and young adults with a medium level of education. Although they lived closer to the project, they were less likely to frequently use the train line under discussion.

This cluster also exhibited neutral to positive evaluations. Demographically, these respondents were mostly middle-aged to seniors, with a medium to high educational background. Members of this cluster were more likely to be infrequent users of both the train and the specific line in discussion.

The third largest cluster displayed some variation across the consultations. In the Lelylijn consultation, this cluster was characterized by critical evaluators who were potential users of the line. Conversely, in the Oude Lijn and Mobility Vision consultations, Cluster 3 comprised more neutral evaluators. Demographically, Cluster 3 included middle-aged individuals with medium education levels, with varying proximity to the project and train usage patterns across consultations.

The smallest cluster consistently represented the most critical evaluators. In both the Lelylijn and Oude Lijn consultations, these respondents were non-users of the line. In the Mobility Vision consultation, this cluster similarly comprised critical evaluators, including middle-aged and young adults, predominantly men, with higher education levels and infrequent train use.

The comparisons of LCCA results between PVE consultations indicate that there are consistent segments that have emerged across all three consultations. The two most positive groups consist of middle-aged individuals and young adults with a moderate to high level of education. The cluster with the highest positive evaluation tends to consist of individuals who live in close proximity to the project and occasionally use the train. In contrast, the neutral-positive

Table 4: Crosstabulation Analysis Results: Sociodemographic Factors vs. Face Validity Evaluations. NC = Not Checked; NS = Not Significant.

Statement	Consultation	Age *Older*	Gender *Women*	Education *High Ed*	Work *Full-time*	Proximity *Proximate*	Freq. Train *Frequent*	Freq. Line *Frequent*
Steering choices	Lelylijn	(+)	(-)	(-)	(-)	(+)	(-)	NS
Difficult to understand	Oude Lijn	NS	NS	(-)	(-)	(-)	(-)	(-)
	Mob. Vision	(-)	NS	(-)	(+)	NC	NC	NC
Fair investigation	Lelylijn	(-)	NS	(-)	(-)	(+)	(+)	(+)
	Oude Lijn	(-)	NS	NS	NS	(+)	(+)	(+)
	Mob. Vision	(+)	NS	(+)	NS	NC	NC	NC
Important topic	Lelylijn	(-)	(+)	NS	NS	(+)	(+)	(+)
	Oude Lijn	NS	(-)	(+)	(+)	(+)	(+)	(+)
	Mob. Vision	(+)	(-)	(+)	NS	NC	NC	NC
Learn government choices	Lelylijn	(-)	(+)	NS	(-)	NS	(+)	(+)
	Oude Lijn	(-)	(+)	NS	NS	NS	(+)	NS
	Mob. Vision	NS	(+)	(-)	NS	NC	NC	NC
Used more	Lelylijn	(-)	(+)	(-)	NS	NS	(+)	(+)
	Oude Lijn	NS	NS	(+)	(+)	(+)	(+)	(+)
	Mob. Vision	(+)	NS	(+)	NS	NC	NC	NC
Increase decision's acceptance	Lelylijn	(-)	(+)	(-)	NS	NS	(+)	(+)
	Oude Lijn	NS	NS	(+)	(+)	(+)	(+)	(+)
	Mob. Vision	(+)	NS	(+)	NS	NC	NC	NC
Increase trusts in government	Lelylijn	(-)	(-)	(-)	NS	NS	(+)	(+)
	Oude Lijn	NS	NS	NS	NS	NS	NS	NS
	Mob. Vision	NS	NS	(-)	NS	NC	NC	NC

Table 5: Comparison of Clusters between Consultations: Highest Probability Level per Cluster. NC = Not Checked.

	Lelylijn				**Oude Lijn**				**Mobility Vision**			
	Cluster1	**Cluster2**	**Cluster3**	**Cluster4**	**Cluster1**	**Cluster2**	**Cluster3**	**Cluster4**	**Cluster1**	**Cluster2**	**Cluster3**	**Cluster4**
Percentage	54%	24%	11%	10%	48%	21%	18%	13%	50%	30%	15%	5%
Indicators												
(+) Statement	7	5	5	2	7	6	4	-	7	4	3	2
(-) Statement	-	-	1	3	-	-	-	-	-	-	-	3
Neutral	-	2	1	2	-	1	3	7	-	3	4	2
Avg Ratings	7.93	7.51	6.08	5.58	8.19	7.47	7.50	6.52	8.09	7.45	6.46	5.93
Covariates												
Age	35 to 65	35 to 65	35 to 65	35 to 65	35 to 65	35 to 65	35 to 65	35 to 65	35 to 65	35 to 65	35 to 65	35 to 65
Gender	Man	Woman	Man	Man	Man	Woman	Man	Woman	Man	Man	Woman	Man
Education	Medium	High	Medium	High	Medium	Medium	High	Low	Medium	High	Low	High
Proximity	Yes	Yes	No	No	Yes	No	Yes	No	NC	NC	NC	NC
Freq Train	Rarely	Rarely	Rarely	Rarely	Frequently	Rarely	Frequently	Rarely	NC	NC	NC	NC
Freq Line	Frequently	Rarely	Frequently	Rarely	Rarely	Rarely	Frequently	Rarely	NC	NC	NC	NC

group, seldom make use of the train and plan to use the line. The two least positive clusters often comprise individuals ranging from young adults to middle-aged persons. Individuals belonging to these groups are more prone to possessing a lower level of educational level compared to the first two positive groups.

5) Multinomial Logistic Regression (MLR):

Multinomial Logistic Regression (MLR) was conducted for each PVE consultation to examine the influence of various respondent characteristics on their assessments of face validity statements. This analysis helps identify which demographic factors are most strongly associated with particular perceptions of the consultation process. The MLR analysis revealed several significant effects of respondents' characteristics on their evaluations of face validity. While some trends were consistent across all consultations, others varied.

Table 6 summarizes the likelihood of different demographic groups agreeing with face validity statements compared to a benchmark group. Consider the statement "Steering choices" in the Lelylijn consultation. The table shows that middle-aged participants (35-65 years) are marked with "A," meaning they are more likely to agree that the consultation steers their choices, compared to seniors (65+ years). Similarly, the table indicates that individuals with high education levels are more likely to disagree ("ND") with this statement than those with lower education.

Older individuals (aged 65 and above) generally rated the consultation process more positively. They were more likely to perceive the consultation as fair, find it easier to navigate, and recognize the importance of the topics discussed. However, the correlation between age and opinions on whether the consultation provided new knowledge or increased trust in the government was less consistent.

Table 6: Results of Multinomial Logistic Regression Analysis on Various Statements, Demographic Factors, and Usage Frequencies. NS = Non-significant. A = Agree. N = Neutral. NA = Neutral-Agree. ND = Neutral-Disagree. D = Disagree. NC = Not Checked. DA = Disagree-Agree.

Statement	**Consultation**	**Age**		**Gender**	**Education**		**Occupation**		**Proximity**	**Freq Use Train**		**Freq Use Line**	
Level vs		<35	35-65	Men	Low	Medium	Unemployed	Part-time	No	Low	Mid	Low	Mid
Benchmark Level		*>65*		*Women*	*High*		*Full-time*		*Proximate*	*High*		*High*	
Steering choices	Lelylijn	NS	A	NA	N	ND	NA	NS	D	NA	N	ND	ND
Difficult to understand	Oudelijn	NA	A	NS	NA	A	NA	NS	A	D	NC	NS	NC
	Mob. Vision	A	A	DA	NA	NA	D	NC	NC	NC	NC	NC	NC
Fair investigation	Lelylijn	A	NS	NS	DA	NS	A	NS	N	ND	D	NS	NS
	Oudelijn	NS	D	D	NS	NS	NS	NS	NS	D	NC	NS	NS
	Mob. Vision	D	D	D	ND	N	NS	NC	NC	NC	NC	NC	NC
Important topic	Lelylijn	DA	NS	ND	A	DA	NS	NS	ND	NS	NS	ND	D
	Oudelijn	ND	ND	NS	N	NS	N	N	ND	D	NC	ND	NS
	Mob. Vision	D	D	DA	ND	ND	NS	NC	NC	NC	NC	NC	NC
Learn government choices	Lelylijn	A	NA	N	NS	A	NS	NS	D	N	DA	NS	D
	Oudelijn	NS	NS	D	NA	NA	NS	NS	A	N	NC	DA	NC
	Mob. Vision	DA	D	D	NA	NA	NS	NC	NC	NC	NC	NC	NC
Used more	Lelylijn	DA	NS	ND	A	A	N	NS	NS	N	NS	NS	NS
	Oudelijn	D	D	D	NS	NA	NS	NS	ND	D	NC	D	NC
	Mob. Vision	NS	A	DA	ND	NS	A	NC	NC	NC	NC	NC	NC
Increase decision's acceptance	Lelylijn	DA	DA	A	DA	NA	D	NS	NA	D	A	D	ND
	Oudelijn	D	NS	D	A	A	NS	NS	NS	NS	NC	D	NC
	Mob. Vision	D	D	DA	NA	A	NS	NC	NC	NC	NC	NC	NC
Increase trusts in government	Lelylijn	DA	DA	DA	NA	NA	NS	NS	NS	D	D	D	NS
	Oudelijn	D	NS	D	NS	NA	NS	NS	NS	D	NC	NS	NC
	Mob. Vision	NS	NS	DA	NA	A	NS	NC	NC	NC	NC	NC	NC

Men consistently exhibited lower levels of perceived fairness, greater difficulty with the consultation process, and lower perceived information acquisition. They were also less likely to support the continued use of this method or trust in the government's decisions compared to women. However, gender had an inconsistent effect on the perceived importance of the consultation topics.

Higher educational attainment was associated with lower likelihoods of perceiving the consultation as a source of new knowledge or as increasing trust in the government. Conversely, more educated participants generally found the consultation process less difficult. The impact of education on perceptions of fairness and importance of the consultation varied across consultations.

Full-time workers were generally less likely to view the consultation as honest, though this trend was not strongly consistent. There was some evidence that full-time employees found the consultation process less challenging, but again, this was not uniformly observed.

Respondents living farther from the project were more likely to perceive the consultation as difficult. However, proximity was consistently associated with viewing the consultation as important. The impact of proximity on the perception of gaining knowledge through the consultation was uncertain.

Frequent train users consistently perceived the consultation as an honest investigation and expressed higher confidence in the government's decisions. They were also less likely to find the consultation process difficult or directive, and more likely to view the topics discussed as important. However, the correlation between train use frequency and perceived knowledge gain from the consultation was inconclusive.

Individuals who planned to use the train line more frequently were more likely to perceive the consultation as steering their decisions. Despite this, these respondents also showed a greater likelihood of viewing the consultation topics as important and being more accepting of the final decisions. The effect of planned line use on perceived knowledge gain was less clear.

B. Underlying Reasons for Negative Evaluations on PVE Consultations

The content analysis of respondents' negative feedback on PVE consultations is summarized in Table 7. The analysis focused on 2,084 filtered responses from the original dataset of 4,691 respondents, excluding empty, unrelated, or positive comments. The feedback was categorized into five dimensions of face validity: feasibility, completeness, acceptance, transparency, and clarity.

The feasibility dimension captured the most feedback, with a notable focus on the Discrete Choice Experiment (DCE) component of the Lelylijn consultation. The most common complaints were about confusing choice experiments (17%) and the consultation being too long (8%). Additionally, 55 respondents highlighted a poor overall experience, often citing issues like difficulty using mobile devices to complete the survey and the inability to navigate back to previous pages.

In terms of completeness dimension, Respondents frequently expressed concerns about the lack of information necessary to make informed decisions. The most common issues included insufficient information on what would happen after the consultation (13%) and limited details about the project itself (10%). Many participants felt that they were not given enough context to understand the project's potential impact.

Table 7: Negative Comments in Lelylijn Consultation and their Frequency of Mentions

Dimension	Category	# of Mentions (%)
Feasibility	Confusing choice experiments	347 (17%)
	Consultation is too long	173 (8%)
	Points allocation is confusing	67 (3%)
	Bad consultation experience	55 (3%)
	Too many/ little number of questions	47 (2%)
	Too complex	26 (1%)
	Only in Dutch	9 (0%)
Completeness	Limited info on consultation follow up	272 (13%)
	Limited info about the project	200 (10%)
	Limited info about the project's impact	53 (3%)
	Limited info about what experts' think	15 (1%)
Acceptance	Online form of consultation	165 (8%)
	Questions lack of depth	156 (7%)
	Limited chance of giving opinion	154 (7%)
	Government/ Experts must take action	18 (1%)
	Focused on Randstad / Northern area only	9 (0%)
	Preference to invest in other projects	5 (0%)
Transparency	Steering into certain direction	188 (9%)
	Others/ Trust Issue	27 (1%)
	Security is questionable	9 (0%)
	Anonymous	2 (0%)
Clarity	Unclear in general	54 (3%)
	Unclear questions/ choices	33 (2%)

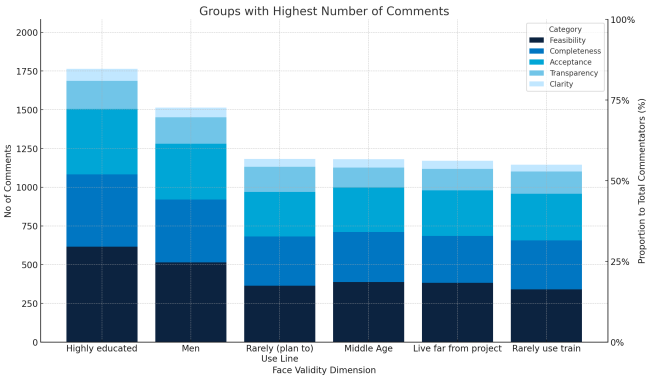


Fig. 4: Groups with Highest Number of Comments per Face Validity Dimension

Regarding the acceptance dimension, comments about the online format of the consultation were prominent, with 8% of the total comments were shown to be concerning it. There were also criticisms about the depth of the questions (7%) and the limited opportunities to provide their opinions (7%). These issues suggest that some participants felt the consultation was not a genuine avenue for meaningful input.

The transparency dimension raised issues about trust and the perception that the consultation was steering respondents towards a predetermined outcome. Nearly 9% of the comments were about the consultation was designed to favor the implementation of the Lelylijn, with several expressing skepticism about the government's intentions. Concerns about the security and anonymity of the consultation were also noted, though less frequently.

A smaller proportion of comments (3%) were about the consultation being unclear, with some pointing specifically to unclear questions or choices (2%). These issues highlight the need for clearer communication in the design of future consultations.

Comments Given numerous times by "Negative" Evaluators Group

The quantitative analysis in Section V-A identified several demographic groups that consistently rated the face validity of PVE consultations negatively. These groups include young adults, men, individuals with high education levels, people living far from the project, and infrequent train users. The following section elaborates on the specific issues raised by these groups:

This group submitted 1,764 of the 2,084 negative comments. Their most frequent complaint was the confusing nature of the choice experiments (296 mentions), which contradicts the quantitative finding that individuals with lower education levels found the consultation more difficult. This suggests that while highly educated respondents may not find the overall consultation difficult, they are critical of specific elements, such as the choice experiments. Other common issues included insufficient information during the consultation (231 mentions) and a lack of project details (174 mentions), potentially explaining why this group reported lower levels of perceived knowledge acquisition and trust.

With 1,514 unfavorable comments given during the PVE consultation, men are the second most frequent critics. Similar to highly educated individuals, men provided the most comments on confusing choice experiments (255 mentions), insufficient information on the follow-up of consultation (197 mentions), and inadequate information about the project (157 mentions). This finding supports the quantitative analysis results that indicate that men are less inclined to perceive this consultation as being honest, learn knowledge about the government's decisions, and exhibit acceptance and confidence in the government's choices. Additionally, they are more inclined to see the consultation as difficult. In addition, the fourth most frequently stated comment, with 147 mentions, also aligns with the quantitative analysis findings indicating men are more likely to believe that this consultation is steering into a certain direction.

The third largest demographic that expressed negative feedback consists of individuals who infrequently (intend to) use Lelylijn. The quantitative research revealed that this group has less of a tendency to consider this topic to be important for expressing their viewpoint and to accept the government's decisions. Although the majority of comments in this group fell into the same categories as the two groups outlined before, there were explicit comments made by individuals who are not potential users of Lelylijn. These comments may offer an explanation for the poor ratings. Quoting a resident, *"I live in Eindhoven, so my opinion is not really of great importance in terms of participation in this decision-making"*. These individuals reside far from Lelylijn and hence have a lower likelihood of intending to utilize Lelylijn. Consequently, they believe that the subject lacks importance, thus making it unnecessary for them to express their opinion on.

One notable category in which this group surpasses the

preceding two major groups is the matter of trust issues, with a total of 21 mentions. *"I have the feeling that the decision about the Lely line has already been made. This without proper consultation with the people who are most affected by it,"* as stated by an individual. In addition, with regards to the minimum information on follow-up of this consultation, someone has commented that it is not completely evident how this consultation would be taken into account in the decision-making process. This could be attributed to the less accepted the choices made by the government by this group.

The quantitative research reveals that young adults have a lower tendency to perceive PVE consultation as being important. Additionally, they are more inclined to perceive this consultation as difficult and less inclined to accept the government's conclusions. Nevertheless, although young adults had a low assessment of the face validity of PVE consultations, it is noteworthy that the majority of comments in the Lelylijn consultation were provided by middle-aged individuals (1,180 comments), while young adults contributed just 517 comments. This suggests that young individuals are less inclined to express their opinions regarding the consultation compared to middle-aged individuals. Another potential cause of this low number of feedback may be that young adults perceive this topic as less important, leading them to neglect providing detailed comments during the consultation.

Regarding the category of comments, the most expressed comment was about the confusing choice experiment (96 mentions), while the complaint about the confusing allocation of points received 33 mentions. These comments shed light on why this group found the consultation difficult to understand. Furthermore, there were multiple references to the lacking information regarding the follow-up of the consultation (57) and the project (57), which may also result in lack of acceptance of the government's decisions.

Individuals residing far from Lelylijn reported a lower perception of the consultation's importance. Many expressed doubts about the relevance of the consultation to their lives, particularly due to the online format, which allowed participation from those not directly affected by the project. Comments such as *"Opinions of residents who will live close to the line are of the utmost importance"* reflect this sentiment.

Infrequent train users exhibit a negative perspective in multiple statements in PVE consultations. They are less likely to perceive the consultation as honest, less inclined to perceive the topic to be important, and less inclined to accept and have confidence in the government's decisions. This phenomenon is evident in the substantial amount of comments relating to the insufficient details regarding the follow-up of the consultation (177) and the fact that the consultation is being guided towards a specific direction (114). One comment that accurately summarizes the above judgment is, *"The questions are very guiding. The entire setting of the consultation already assumes that the line will be built anyway. Can't really be called objective. It is also poorly publicized that this consultation is taking place, probably with the part that mainly proponents will complete it."*

Another comment was made about the project's lack of comprehensive information (97 mentions), stating *"Too little attention to the negative sides of this train connection. Consider financial risk, landscape impact, etc"*. The statement indicates that individuals who do not utilize trains may be critical about the adverse consequences associated with the construction of a new rail line. This could also account for the negative assessment of the PVE in the aforementioned categories.

C. Experts Perspective Exploration

The discussion with PVE expperts began with a presentation of the quantitative and qualitative findings from Section V-A and Section V-B, followed by a Q&A session to clarify any points. The central question posed to the group was: *"How can we address the negative evaluations provided by certain groups?"* The discussion was open-ended, allowing participants to freely express their opinions. The primary topics discussed are summarized as follows:

1) Provision of Additional Information
   The availability of sufficient information is critical in consultation methods like PVE. Quantitative analysis in Section V-A revealed that approximately 70% of participants felt they had learned about the government's choices, while 10% felt otherwise. Content analysis in section V-B highlighted that many respondents (340 mentions) felt the consultation lacked adequate information, particularly regarding the project itself, its impacts, expert opinions, and follow-up actions.
   The discussion began with a suggestion to include hyperlinks to background information. However, this was not recommended, as it might distract participants from the consultation. An alternative was to use a button linking to methodological background information. Despite this, there was skepticism within the group about providing more information, as it might lead to increased doubt among participants, reflecting findings from studies like [63], which suggest that too much information can generate more questions. The consensus was that while additional information could be valuable, it must be carefully curated to clarify rather than complicate the consultation process.
2) Offering Concise and Extensive Versions
   Although overall ratings for the PVE consultations were positive, approximately 9.7% of respondents found the consultation difficult to understand. Content analysis revealed that the feasibility dimension—how easy it was to answer questions or choose options—was a common source of dissatisfaction, with 724 mentions of issues ranging from confusing questions to the consultation being too lengthy or complex.
   A suggestion from the PVE consultation was to provide both concise and extensive versions of the consultation, allowing participants to choose based on their comfort level. However, the group acknowledged the risks of this approach, particularly concerning how it might

affect those who already view the consultation positively. The consensus leaned towards providing personalized information, tailored to the needs of different participants, while maintaining simplicity where possible. This approach would involve understanding the specific needs of participants and offering relevant information accordingly.

3) Randomization of Options
   The statement *"If many people participate in this consultation, it will increase my acceptance of the final decisions on this topic"* received the lowest average rating (64.8%) compared to other face validity statements. Content analysis also indicated concerns regarding the acceptance and transparency dimensions. One PVE respondent noted that the order of options did not seem natural, suggesting that *"option 1 usually has the most points in your analysis because anchoring takes place here."*
   The group discussion revealed that while options are randomized in each consultation, participants are often unaware of this. It was suggested that a button could be added to inform participants that the options are randomized. However, concerns were raised about how to incorporate this information without overloading participants, particularly in the digital version of the consultation. There was also concern that some participants might feel disadvantaged if they perceive others as having more access to information.
4) Increasing Legitimacy and Trust
   Results from Section V-A indicated that 15.9% of respondents felt the consultation was steering them in a particular direction. This perception may relate to the numerous comments on acceptance and transparency dimensions identified in the content analysis.
   One proposed solution was to explain the inherent steering nature of PVE, as the method limits options and solutions by design to focus on what is most relevant for policymakers. Suggestions included explaining this through animations or instructional videos and adding an option for participants to indicate if they feel an option is missing, followed by an explanation of why certain choices were included. This approach requires careful consideration of potential questions participants may have during the process.
   Another suggestion was to address legitimate perspectives that are not part of the choice design, such as those held by individuals who do not believe in climate change but are still asked to choose measures to combat it. The group discussed the importance of considering these perspectives in the design phase to ensure inclusiveness and legitimacy.
   Finally, the group explored the idea of "name-dropping" trustworthy institutions involved in the consultation process to enhance legitimacy. By highlighting collaborations with recognized organizations, the consultation could potentially gain more trust from participants. However, this approach also requires consideration of how these institutions might respond if questioned by the public about their involvement.

VI. Discussion and Recommendation

This study aimed to understand better the characteristics of respondent that (in)consistently evaluate Public Value Evaluation (PVE) consultations in certain way, specifically in within transportation PVE. The following discussion solidifies the findings, acknowledges the limitations, and explores the broader implications for public participation methods.

The concept of face validity, as discussed in Chapter Section III, is crucial in public participation methods, including PVE. It is used to assess how well a method captures respondents' opinions on policy options, particularly in terms of clarity, relevance, and comprehensiveness. However, the reliance on face validity alone might be insufficient for a robust evaluation of public participation methods. As [48] suggest, employing multiple validation approaches, such as content and criterion validity, is necessary to achieve a comprehensive assessment. Therefore, future research should explore these additional forms of validity to provide a more thorough evaluation of PVE consultations.

The discussion then addresses the relationship between respondents' characteristics and their evaluations of PVE consultations. This study utilized weighted datasets to align the sample characteristics with the general population in the Netherlands, enhancing the representativeness of the findings. However, the process of weighting introduced potential concerns, such as simplifications in the analysis that could affect the outcomes. Future research should aim for consistency in demographic characteristics across consultations to minimize discrepancies and improve the robustness of the conclusions.

Another significant discussion point involves the dataset used for analysis. During data preparation, responses such as "I don't know" were excluded to maintain consistency across analyses, which led to a reduction in the number of data points. Despite this, the datasets remained substantial, with approximately 4,000 respondents each. The high overall ratings for the consultations observed in Section V-A are notable, as such positive evaluations are rare in past PVE consultations. This trend may be attributed to the nature of the consultations, which predominantly presented favorable scenarios. However, the exclusion of uncertain responses could have influenced these ratings. Future studies should investigate whether excluding certain responses significantly impacts overall ratings and should consider including follow-up questions to understand why participants might drop out of the consultation process.

The perception of PVE consultations as steering or directive is another critical discussion point. In the Lelylijn consultation, a notable percentage of respondents perceived the consultation as influencing their choices, with many expressing neutral responses. This finding aligns with previous research, but it also raises concerns about the transparency of the consultation process. Experts suggest that the inherently directive nature of

PVE should be clearly communicated to respondents to enhance transparency and trust. Future PVE consultations might benefit from exploring whether framing the consultation in a more positive light, as seen in [3] study, affects participants' perceptions of objectivity.

The discussion also highlights the challenges associated with assessing the feasibility of PVE consultations. Approximately 10% of respondents in the Oude Lijn and Mobility Vision datasets found the consultation difficult to understand. Content analysis from the Lelylijn consultation revealed that concerns about the complexity of the consultation were often voiced by highly educated individuals, who speculated that less educated or digitally literate respondents might struggle in PVE participation. This finding suggests the need for further research to analyze the demographics of respondents who find the consultation difficult, as well as conducting similar analyses on the other datasets to gain a comprehensive understanding of the issue.

Digital literacy also emerged as a potential source of bias in PVE consultations. The use of a panel of respondents, who are typically more digitally literate, may result in an inflated assessment of face validity. To mitigate this bias and ensure a more inclusive consultation process, future PVE designs could consider reaching out to participants through traditional mail or linking the consultation to citizens' IDs. However, any approach must carefully balance inclusiveness with ethical considerations, such as maintaining respondent anonymity.

The analysis of respondents' characteristics revealed that some factors, such as age and proximity, consistently influenced evaluations of PVE consultations. For instance, young adults tended to rate PVE consultations highly overall, but their evaluations were less positive compared to older respondents. This finding aligns with previous research by [1], although discrepancies were observed in how proximity was defined and its impact on evaluations. The study suggests that a broader definition of proximity may yield different results than more localized studies. Future research should consider not only where respondents live but also how close their daily activities are to the project in question, providing a more nuanced understanding of the influence of proximity on face validity evaluations.

Lastly, the discussion addressed the content analysis of negative feedback from the Lelylijn consultation. Although the content analysis was limited to the Lelylijn dataset, it is evident that several concerns identified in that consultation were already addressed in the subsequent Oude Lijn and Mobility Vision consultations. This may explain the higher overall ratings observed in those consultations compared to Lelylijn. Conducting content analyses for both the Oude Lijn and Mobility Vision datasets would provide valuable insights and help confirm this observation. However, it is important to note that content analysis is inherently qualitative and subjective. To enhance the consistency and robustness of future research, incorporating peer review into the content analysis process would be beneficial.

Reflection on PVE and Broad Welfare (BW) Criteria

This research is closely aligned with the development of forward-thinking and inclusive transport designs, consistent with the broad welfare perspective. PVE methodology has been recognized as well-suited for evaluating public policy based on broad prosperity, as it encourages public participation, captures diverse demographic perspectives, and ensures that all essential values and concerns are considered. In addition to assessing the face validity of the PVE method, this study explored whether PVE consultations correspond with welfare beyond GDP perspective, or broad welfare, also known as *"brede welvaart"* in the Dutch context.

The Broad Welfare criteria, developed by TNO, include considerations such as time periods, geographic scope, participation and co-creation, and the distribution of benefits and burdens. PVE consultations meet these criteria by allowing for flexible customization, considering specific geographic locations, enabling participation and co-creation, and analyzing the distribution of impacts across different demographic groups. Although PVE does not directly integrate different government silos, it encourages a comprehensive approach by involving multiple ministries and stakeholders in the design phase. Additionally, PVE consultations engage participants in assessing trade-offs and synergies, ensuring that policy decisions are informed by a broad and inclusive perspective.

While PVE is utilized as a transportation planning tool within the broad welfare perspective, it is important to recognize that this study cannot be directly compared to evaluations of tools using BW criteria. This study focuses on the face validity of PVE, specifically how users or respondents perceive the tool in terms of its effectiveness in measuring what it intends to measure. In contrast, BW criteria are employed by experts or organizations to assess various transportation planning tools. Nevertheless, there are notable overlaps and connections between this study and BW criteria, suggesting that PVE is indeed aligned with the principles of broad welfare. Further research could be valuable in exploring this relationship more comprehensively.

Recommendations for Future Research

Future research should delve deeper into the segmentation of PVE consultation evaluations to uncover patterns that may be obscured by broader categorizations. This study found that characteristics such as financial status, type of home, and area of living showed varying associations across different consultations. More granular categorization of these attributes could reveal new associations and provide a nuanced understanding of their influence on respondents' evaluations. Additionally, further investigation into other types of validity beyond face validity, such as content and criterion validity, is essential to enhance the robustness of PVE as a public participation tool. Future studies should also examine the outcomes of content analysis in all PVE consultations and consider peer-reviewed analysis to ensure consistency and robustness.

Recommendations for Future PVE Consultations

To improve the effectiveness and face validity of future PVE consultations, several key strategies emerged from the expert discussions. Providing personalized information while maintaining simplicity is crucial to catering to different participant preferences without overwhelming them. Clarifying the randomization of options and the inherently steering nature of PVE can enhance transparency and trust among participants. Involving trusted institutions and clearly explaining the rationale behind the options can further legitimize the consultation process. Addressing the negative evaluations through these strategies will help refine PVE consultations, making them more inclusive, transparent, and effective in capturing public preferences for transportation projects.

VII. Conclusions

This study provided critical insights into the Participatory Value Evaluation (PVE) method, particularly in its application to transportation planning. As PVE is increasingly used as both a public participation tool and an evaluation method, this research aimed to address the limited empirical evidence on its validity and perception across different population segments. The study was guided by three key research questions, which are summarized below alongside the conclusions drawn from the analysis.

A. Distinct Population Segments and Face Validity Evaluations

The analysis revealed that certain demographic groups, including young adults, men, highly educated individuals, renters, and frequent train users, generally rated PVE consultations more positively. However, there were notable variations in how these groups assessed specific aspects of face validity. Four distinct respondent clusters were identified, ranging from highly positive to negative evaluators. Positive evaluators typically comprised younger adults with medium to high education levels, who lived closer to the project and were frequent users of relevant transportation infrastructure. In contrast, negative evaluators were more likely to be middle-aged, with varied education levels, and lived farther from the projects, often using the project or line less frequently.

B. Consistency of Face Validity Evaluations Across PVE Experiments

While some demographic characteristics, such as being young or male, consistently influenced higher overall ratings, inconsistencies were found in other areas. For instance, men and highly educated individuals, though generally positive overall, were more critical when assessing whether the consultation increased their knowledge or trust in government decisions. Multinomial logistic regression analyses confirmed that while young adults and frequent train users were more likely to rate the consultations positively overall, they were also more discerning in their evaluations of specific face validity statements. This suggests that the consistency of evaluations varies depending on both demographic factors and the specific context of the consultation.

C. Factors Contributing to Negative Assessments of PVE

Content analysis of feedback from the Lelylijn consultation highlighted the feasibility dimension as the most criticized aspect, particularly regarding the complexity and length of the consultation process. Concerns about the completeness of information provided, as well as transparency and acceptance, were also prevalent. Interestingly, highly educated individuals and men, who generally rated the consultations positively, were also among the most vocal critics, particularly regarding the lack of information and the perceived steering nature of the consultation. This finding emphasizes the importance of addressing these concerns to improve the overall perception of PVE.

D. Strategies for Improving PVE Face Validity

Several strategies were identified to enhance the face validity of PVE consultations. Simplifying the consultation process and offering both concise and extensive versions can cater to diverse participant needs, enhancing overall understanding, especially for those with lower educational levels. Additionally, providing information in a way that clarifies rather than overwhelms participants is crucial, as overloading them with details might lead to confusion. To address concerns about transparency and perceived steering, it is recommended to explicitly communicate the rationale behind option selection and involve reputable institutions in the consultation process to build trust. Moreover, including mechanisms that allow participants to express personal viewpoints not covered by existing options can foster a sense of inclusiveness and ensure that diverse perspectives are acknowledged. These strategies collectively aim to enhance the credibility, effectiveness, and participant engagement in PVE consultations, ensuring they remain a robust tool for public involvement in policy-making.

E. Evaluation of Hypotheses

The evaluation of the hypotheses presented in this study reveals a nuanced relationship between demographic factors and the perceived face validity of Public Value Evaluations (PVE) in transportation planning. The findings indicate that older individuals, while giving lower overall ratings compared to younger adults, were more critical of specific aspects such as honesty and effectiveness in increasing the acceptance of government decisions, partially supporting the hypothesis. Similarly, although the hypothesis posited that women would generally have a more positive attitude towards PVE, it was found that men rated consultations higher overall. However, women were more positive about certain aspects of face validity, suggesting a mixed support for this hypothesis.

The study also found strong support for the hypothesis that highly educated respondents view PVE positively but with a critical perspective, particularly in areas related to learning from the consultation and trust in government decisions. Frequent users of the train or the specific project under consideration, as well as individuals living closer to the project, tended to evaluate PVE more positively, with the latter group particularly emphasizing the importance and honesty

of the consultation. These findings underscore the importance of considering demographic differences when evaluating the effectiveness of PVE consultations and suggest that future research should further explore these dynamics across different contexts and populations. This approach will help ensure that PVE remains an inclusive and effective tool for public participation in policy-making.

REFERENCES

[1] C. Tuit, "Face Validity of the Participatory Value Evaluation Method: A Case Study of the Schiphol Environmental Council," M.Sc. thesis, Delft University of Technology, Delft, Netherlands, 2022. Unpublished master's thesis.

[2] H. Taherdoost, "Validity and Reliability of the Research Instrument; How to Test the Validation of a Questionnaire/Survey in a Research," Aug. 10, 2016.

[3] E. Golan, "Step Into the Driver's Seat: A Participatory Value Evaluation of the Public Transport Policy Preferences of the Tel Aviv Metropolitan Area & Israeli Face Validity Analysis," M.Sc. thesis, Delft University of Technology, Delft, Netherlands, Aug. 2023.

[4] J. Gaber and S. L. Gaber, "Using face validity to recognize empirical community observations," *Evaluation and Program Planning*, vol. 33, no. 2, pp. 138–146, 2010. doi: 10.1016/j.evalprogplan.2009.08.001.

[5] A. Ramadani, "Face Validity in Participatory Value Evaluation: Exploring Segment-specific Perceptions and their Influence on Transport Decision-Making," M.Sc. thesis, Delft University of Technology, Delft, Netherlands, 2024, unpublished.

[6] L. Delaney, "Descriptive statistics: simply telling a story," *African Journal of Midwifery and Women's Health*, vol. 4, no. 1, 2010. doi: 10.12968/ajmw.2010.4.1.46313.

[7] J. Liang, G. Bi, and C. Zhan, "Multinomial and ordinal Logistic regression analyses with multi-categorical variables using R," *Annals of Translational Medicine*, vol. 8, no. 16, p. 982, 2020. doi: 10.21037/atm-2020-57.

[8] H. Aacharya, "Content Analysis," *EPRA International Journal of Research and Development (IJRD)*, vol. 7, no. 4, pp. 177–180, 2022. doi: 10.36713/epra2016.

[9] Columbia University, "Content Analysis," 2023. [Online]. Available: https://www.publichealth.columbia.edu/research/population-health-methods/content-analysis. [Accessed: 2-Jul-2024].

[10] A. Morris, A. J. Onwuegbuzie, and H. R. Gerber, "Using Expert Interviews within MODES in Online and Offline Spaces to Extend Comprehensive Literature Review Processes," *The Qualitative Report*, vol. 23, no. 8, pp. 1777–1798, 2018. doi: 10.46743/2160-3715/2018.3039.

[11] P. N. Ghauri, K. Gronhaug, and R. Strange, *Research Methods in Business Studies*. Cambridge: Cambridge University Press, 2020.

[12] A. de Ruijter, *[Unpublished Article]*, 2022.

[13] P. Burton, "Conceptual, Theoretical and Practical Issues in Measuring the Benefits of Public Participation," *Evaluation*, vol. 15, no. 3, pp. 263–284, 2009. doi: 10.1177/1356389009105881.

[14] A. de Ruijter, *[Unpublished Article]*, 2023.

[15] I. Dalawi, M. Isa, X. W. Chen, Z. Azhar, and N. Aimran, "Development of the Malay Language - content validity & face validity analysis of MUAPHQ C-19," *BMC Public Health*, vol. 23, Art. no. 16044-5, 2023. doi: 10.1186/s12889-023-16044-5.

[16] K. Glenk, J. Meyerhoff, S. Colombo, and M. Faccioli, "Enhancing the face validity of choice experiments: A simple diagnostic check," *Ecological Economics*, vol. 221, p. 108160, 2024. doi: 10.1016/j.ecolecon.2024.108160.

[17] O. Babalola, J. Goudge, J. Levin, C. Brown, and F. Griffiths, "Assessing the Utility of a Quality-of-Care Assessment Tool Used in Assessing Comprehensive Care Services Provided by Community Health Workers in South Africa," *Frontiers in Public Health*, vol. 10, p. 868252, 2022, doi: 10.3389/fpubh.2022.868252.

[18] I. Dalawi, M. Isa, X. W. Chen, Z. Azhar, and N. Aimran, "Development of the Malay Language - content validity & face validity analysis of MUAPHQ C-19," *BMC Public Health*, vol. 23, Art. no. 16044-5, 2023. doi: 10.1186/s12889-023-16044-5.

[19] K. Glenk, J. Meyerhoff, S. Colombo, and M. Faccioli, "Enhancing the face validity of choice experiments: A simple diagnostic check," *Ecological Economics*, vol. 221, p. 108160, 2024. doi: 10.1016/j.ecolecon.2024.108160. "

[20] O. Babalola, J. Goudge, J. Levin, C. Brown, and F. Griffiths, ""Assessing the Utility of a Quality-of-Care Assessment Tool Used in Assessing Comprehensive Care Services Provided by Community Health Workers in South Africa,"" *Frontiers in Public Health*, vol. 10, p. 868252, 2022, doi: 10.3389/fpubh.2022.868252."

[21] K. Hermans, N. Spruytte, J. Cohen, C. Van Audenhove, and A. Declercq, "Usefulness, feasibility and face validity of the interRAI Palliative Care instrument according to care professionals in nursing homes: A qualitative study," *International Journal of Nursing Studies*, vol. 62, pp. 90–99, 2016. doi: 10.1016/j.ijnurstu.2016.07.014.

[22] M. A. M. Reffien, A. Ismail, S. Sulong, T. P. Z. T. Bahanuddin, N. M. Salleh, and N. H. Nasir, "Development, Validity and Reliability of an Instrument to Evaluate Implementation Fidelity of the Family Doctor Concept Programme in Public Primary Care Clinics in Malaysia," *Malaysian Family Physician*, vol. 17, no. 1, pp. 20–28, 2022. doi: 10.51866/oa1197. "

[23] J. Carlton, T. Peasgood, C. Mukuria, J. Connell, J. Brazier, K. Ludwig, O. Marten, S. Kreimeier, L. Engel, M. Belizán, Z. Yang, A. Monteiro, M. Kuharic, N. Luo, B. Mulhern, W. Greiner, S. Pickard, and F. Augustovski, ""Generation, Selection, and Face Validation of Items for a New Generic Measure of Quality of Life: The EQ-HWB,"" *Value in Health*, vol. 25, no. 4, pp. 512-524, 2022, doi: 10.1016/j.jval.2021.12.007." "

[24] S. Öztürk, ""Face Validation in Large-Scale Simulation Projects,"" *Mugla Journal of Science and Technology*, vol. 9, no. 1, pp. 24-33, 2023, doi: 10.22531/muglajsci.1179857."

[25] S. Morton, A. Baccini, and A. Morandi, "Content and Face Validity of the Evaluation Tool of Health Information for Consumers (ETHIC)," *International Journal of Environmental Research and Public Health*, vol. 20, no. 6, p. 4872, 2023. doi: 10.3390/ijerph20064872. "

[26] L. Engel, L. Kosowicz, E. Bogatyreva, F. Batchelor, N. Devlin, B. Dow, A. S. Gilbert, B. Mulhern, T. Peasgood, and R. Viney, ""Face Validity of Four Preference-Weighted Quality-of-Life Measures in Residential Aged Care: A Think-Aloud Study,"" *The Patient - Patient-Centered Outcomes Research*, vol. 16, pp. 655-666, 2023, doi: 10.1007/s40271-023-00647-6."

[27] J. C. Banna, N. L. Keim, and M. S. Townsend, "Assessing Face Validity of a Physical Activity Questionnaire for Spanish-Speaking Women in California," *The Journal of Extension*, vol. 49, no. 5, Art. 7, 2023. doi: 10.34068/joe.49.05.07.

[28] N. E. Abdul Hadi, N. A. J. Shafidan, S. R. Razali, Q. M. Abdulateef, and M. S. Abdul Hamid, "The face and content validity of an instrument for measuring financial risk tolerance," *Journal of Computational Innovation and Analytics (JCIA)*, vol. 2, no. 1, pp. 57–88, 2023.

[29] J. Mason *et al.*, "Establishing face and content validity of a survey to assess users' perceptions of automated vehicles," *Journal of Ergonomics*, vol. 8, no. 3, pp. 10–15, 2020. doi: 10.1080/00140139.2020.1800001.

[30] M. Juschten and I. Omann, "Evaluating the relevance, credibility and legitimacy of a novel participatory online tool," *Environmental Science & Policy*, vol. 146, pp. 90–100, 2023. doi: 10.1016/j.envsci.2023.05.001.

[31] B. Nevo, "Face Validity Revisited," *Applied Psychological Measurement*, vol. 9, no. 2, pp. 239–247, 1985. doi: 10.1111/j.1745-3984.1985.tb01065.x.

[32] N. Patel and S. Desai, "ABC of Face Validity for Questionnaire," *International Journal of Pharmaceutical Sciences Review and Research*, vol. 65, pp. 164–168, Nov. 2020. doi: 10.47583/ijpsrr.2020.v65i01.025.

[33] J. Connell, J. Carlton, A. Grundy, *et al.*, "The importance of content and face validity in instrument development: lessons learnt from service users when developing the Recovering Quality of Life measure (ReQoL)," *Quality of Life Research*, vol. 27, pp. 1893–1902, 2018. doi: 10.1007/s11136-018-1847-y.

[34] N. E. Abdul Hadi, N. A. J. Shafidan, S. R. Razali, Q. M. Abdulateef, and M. S. Abdul Hamid, "The face and content validity of an instrument for measuring financial risk tolerance," *Journal of Computational Innovation and Analytics (JCIA)*, vol. 2, no. 1, pp. 57–88, 2023.

[35] B. J. A. Van Wee, D. Annema, and D. Banister, *The Transport System and Transport Policy, An Introduction*. Cheltenham, UK/Northampton, MA, USA: Edward Elgar, 2013.

[36] E. Cascetta and F. Pagliara, "Public engagement for planning and designing transportation system," *Procedia - Social and Behavioral Sciences*, vol. 87, pp. 103–116, 2013.

[37] A. Buijs and F. Boonstra, *Natuurbeleid betwist. Visies op legitimiteit en het natuurbeleid*. KNNV Uitgeverij, 2020.

[38] K. Bickerstaff and G. Walker, "Participatory local governance and transport planning," *Environment and Planning A: Economy and Space*, vol. 33, no. 3, pp. 431–451, 2001, doi: 10.1068/a33173.

[39] W. Ansari and E. Andersson, "Beyond value? Measuring the costs and benefits of public participation," *Journal of Integrated Care*, vol. 19, no. 6, pp. 45–57, 2011.

[40] E. Cascetta, "Transportation systems analysis: Models and applications," in *Proc. Transportation Systems Analysis*, 2009. [Online]. Available: https://api.semanticscholar.org/CorpusID:106443995.

[41] Y. Zheng and H. L. Schachter, "Explaining citizens' e-participation use: The role of perceived advantages," *Public Organization Review*, vol. 17, no. 3, pp. 409–428, 2016, doi: 10.1007/s11115-016-0346-2.

[42] N. Mouter, J. I. Hernandez, and A. V. Itten, "Public participation in crisis policymaking: How 30,000 Dutch citizens advised their government on relaxing COVID-19 lockdown measures," *PLoS ONE*, vol. 16, no. 5, pp. e0250614, 2021, doi: 10.1371/journal.pone.0250614.

[43] S. Spruit and N. Mouter, "1376 inwoners van Súdwest-Fryslân over het toekomstige energiebeleid van hun gemeente: de uitkomsten van een raadpleging," 2020.

[44] J. B. Rosener, "Citizen participation: Can we measure its effectiveness?," *Public Administration Review*, vol. 38, no. 5, pp. 457, 1978, doi: 10.2307/975505.

[45] G. Rowe and L. J. Frewer, "Evaluating public-participation exercises: A research agenda," *Science, Technology, & Human Values*, vol. 29, no. 4, pp. 512–556, 2004, doi: 10.1177/0162243903259197.

[46] J. Gaber and S. L. Gaber, "Using face validity to recognize empirical community observations," *Evaluation and Program Planning*, vol. 33, no. 2, pp. 138–146, 2010, doi: https://doi.org/10.1016/j.evalprogplan.2009.08.001.

[47] S. Boxebeld, N. Mouter, and J. van Exel, "Participatory value evaluation (PVE): A new preference-elicitation method for decision making in healthcare," *Appl Health Econ Health Policy*, vol. 22, pp. 145–154, 2024, doi: https://doi.org/10.1007/s40258-023-00859-9.

[48] K. Bannigan and R. Watson, "Reliability and validity in a nutshell," *Journal of Clinical Nursing*, vol. 18, no. 23, pp. 3237–3243, 2009, doi: 10.1111/j.1365-2702.2009.02939.x.

[49] A. Anastasi and S. Urbina, *Psychological Testing*, 2nd ed., Pearson, 2007.

[50] M. Q. Patton, *Qualitative Research and Evaluation Methods*, 3rd ed., Thousand Oaks, CA: Sage Publications, 2002.

[51] C. Barnett and N. Mahony, "Segmenting publics," unpublished, Nov. 2011.

[52] M. F. Aini, M. Elias, H. Lamers, U. Shariah, P. Brooke, and M. H. Hussin, "Evaluating the usefulness and ease of use of participatory tools for forestry and livelihoods research in Sarawak, Malaysia," *Forests, Trees and Livelihoods*, vol. 26, no. 1, pp. 29-46, 2017, doi: 10.1080/14728028.2016.1246213.

[53] O. Avcıoğlu, "Public participation in local decision-making processes: The case of Bodrum," Master's Thesis, Middle East Technical University, Ankara, Turkey, 2023. [Online].

[54] M. Wester and C. Morn, "Gender and public participation: A study of the Swedish public debate on the Nord Stream gas pipeline from a gender perspective," *Journal of Environmental Assessment Policy and Management*, vol. 15, no. 03, pp. 1350016, 2013, doi: 10.1142/S1464333213500166.

[55] H. Coffé and C. Bolzendahl, "Same game, different rules? Gender differences in political participation," *Social Science Quarterly*, vol. 90, no. 5, pp. 1205–1222, 2009, doi: 10.1111/j.1540-6237.2009.00692.x.

[56] D. C. Glass, H. L. Kelsall, C. Slegers, A. B. Forbes, B. Loff, D. Zion, and L. Fritschi, "A telephone survey of factors affecting willingness to participate in health research surveys," *BMC Public Health*, vol. 15, pp. 1017, 2015, doi: 10.1186/s12889-015-2350-9.

[57] L. A. Onyango, F. N. Mutui, and E. Wabwire, "An analysis of the factors affecting public participation in environmental impact assessment: Case study of selected projects in Nairobi City County, Kenya," *European Scientific Journal*, vol. 15, no. 9, pp. 284-301, 2019, doi: 10.19044/esj.2019.v15n9p284.

[58] V. Carreira, J. R. Machado, and L. Vasconcelos, "Citizens' education level and public participation in environmental and spatial planning public policies: Case study in Lisbon and surrounds counties," *International Journal of Political Science (IJPS)*, vol. 2, no. 3, pp. 25-34, 2016, issn: 2454-9452.

[59] J. Harris, M. Graue, and T. Dunning, "Searching for the impact of participation in health and health research," *BioMed Research International*, vol. 2018, 2018.

[60] N. M. A. Huijts, G. de Vries, and E. J. E. Molin, "A positive shift in the public acceptability of a low-carbon energy project after implementation: The case of a hydrogen fuel station," *Sustainability*, vol. 11, no. 8, Article no. 2220, 2019. [Online]. Available: https://www.mdpi.com/2071-1050/11/8/2220, doi: 10.3390/su11082220.

[61] S. O. Oladeji and others, "Community participation in conservation and management of cultural heritage resources in Yoruba ethnic group of South Western Nigeria," *Journal of Heritage Management*

[62] A. R. Ramadani, "Face validity in participatory value evaluation: Exploring segment-specific perceptions and their influence on transport decision-making," Unpublished master's thesis, Delft University of Technology, Delft, Netherlands, Aug. 2024. [Online].

[63] A. Coulter and J. Ellins, "Effectiveness of strategies for informing, educating, and involving patients," *BMJ*, vol. 335, no. 7609, pp. 24–27, 2007, doi: 10.1136/bmj.39246.581169.80.

B Characteristics of the Sample

This chapter presents a comprehensive overview of the sample characteristics for three consultations: Lelylijn, Oudelijn, and Mobility Vision PVE consultations. The analysis comprises the descriptive analysis of the dataset, conducted after data cleaning. Additionally, it involves the process of assigning weights to factors and the resulting composition of the dataset after factor-weighting. The ultimate dataset was utilized for subsequent quantitative analysis in this study.

B.1 Lelylijn Project

B.1.1 Sample Characteristics before Factor Weighting: Lelylijn

The characteristics of the respondents based on the actual data are presented in Table B.1. The attributes given in the table are pre-coded attributes that have been derived from the available options in the consultation, as demonstrated in Appendix F.

Table B.1: Lelylijn Project Sample Characteristics: Before Factor-Weighting

Respondents' Attribute	Total Respondents*	Respondents %	Population Netherlands %
Gender			
Man	2.332	51%	51%
Woman	2.240	49%	49%
Age			
18-34 years	1.215	27%	27%
35-64 years	2.220	49%	48%
65 years or older	1.137	25%	25%
Education			
Low	1.245	27%	29%
Medium	1.684	37%	37%
High	1.643	36%	35%
Occupation			
Not working	1.773	39%	
Part-time	1.022	22%	
Full-time	1.778	39%	
Financial			
Not enough	263	6%	
Enough	3.191	70%	
More than enough	1.118	24%	
Residency			
Tenant	975	21%	
Lodger	395	9%	
Homeowner	3.202	70%	
Case specific			
Proximity			
Not proximate	2.312	51%	
Proximate	2.261	49%	
Frequency of Using Train			
Rarely	2.869	63%	
Occassionally	925	20%	
(Almost) daily	778	17%	
Frequency of Using Lelylijn			
Rarely	2.650	58%	
Occassionally	1.260	28%	
(Almost) daily	663	15%	

Following the completion of the data-cleaning process, a dataset containing 4.691 respondents was obtained for the Lelylijn consultation, as shown in Table B.1. The demographic profile of Lelylijn consultation respondents differs from the general population of the Netherlands in terms of gender, age, and education, as can be noted. The dataset shows a significant over-representation of men between the ages of 35 and 64 who have a high level of education. An extensive range of statistical studies was con-

ducted on the original dataset. The preliminary findings from the dataset that did not take into account the weight of the data produced less significant and meaningful conclusions. Consequently, the Lelylijn dataset underwent factor weighting (B.1.2) to facilitate subsequent statistical analysis. The utilization of factor weighting led to minor alterations in the proportions of the participants, as demonstrated in Table B.5.

B.1.2 Sample Characteristics after Factor Weighting: Lelylijn

In applying factor weighting to the Lelylijn dataset, three attributes were considered: gender, age, and education level. The composition of respondents (%) was compared to the population proportion of the Netherlands (%) to determine statistical differences using a Chi-square test. The population proportions were referred to as the Expected Proportion (%), while the respondents' proportions were referred to as the Observed Proportion (%).

First, the Chi-square statistic was calculated using formula (B.1). Then, the degrees of freedom (dof) were determined using formula (B.2), and the critical Chi-square value for the given dof was found in a Chi-square distribution table. Subsequently, a weight factor for each attribute level was calculated using formula (B.3).

- Calculate the chi-square statistic using the formula:

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} \tag{B.1}$$

  where O_i is the observed frequency and E_i is the expected frequency for each cell.
- Determine the degrees of freedom for the test, which is calculated as:

$$(\text{rows} - 1) \times (\text{columns} - 1) \tag{B.2}$$

- Look up the critical chi-square value in a chi-square distribution table or use statistical software to find the p-value associated with the chi-square statistic.
- Calculate a weight factor for each attribute level using the formula:

$$\text{Weight Factor} = \frac{E_i}{O_i} \tag{B.3}$$

In the chi-square test, as shown in Table B.2, the Chi-square value for gender (934,88), age (994,53), and education level (3.499,97) are significantly higher than the critical chi-square value for dof 1 and 2. This result indicates significantly different sample than the Netherlands population. Therefore, factor weighting was applied to the data based on these three attributes.

Table B.2: 1st Chi-square Test and Factor-Weighting Lelylijn Dataset

For 1st Iteration				
	Expected Proportion %	Observed Proportion %	Chi-square	Weighting factor
Gender				
Man	51%	74%	410,81	1,00
Woman	49%	26%	427,57	1,00
		Chi-square	838,38	> 3,841
		dof	1	
Age				
18-34 years	27%	25%	3,45	0,81
35-64 years	48%	55%	43,60	0,95
65 years or older	25%	19%	52,13	1,58
		Chi-square	99,18	> 5,991
		dof	2	
Education				
Low	29%	2%	1040,46	2,14
Medium	37%	17%	411,91	1,22
High	35%	81%	2416,69	0,63
		Chi-square	3.869,07	> 5,991
		dof	2	

The results after the first iteration of factor weighting were then re-evaluated to check if the proportions of gender, age, and education level in the dataset were still statistically different from the population. From Table B.3, it was observed that despite the chi-square values of all the three attributes were decreasing, they are still greater than the critical chi-square values. Consequently, a second iteration of factor weighting was applied to the data again.

Table B.3: 2nd Chi-square Test and Factor-Weighting Lelylijn Dataset

For 2nd Iteration				
	Expected Proportion %	Observed Proportion %	Chi-square	Weighting factor
Gender				
Man	51%	58%	41,80	1,00
Woman	49%	42%	43,50	1,00
		Chi-square	85,30	> 3,841
		dof	1	
Age				
18-34 years	27%	26%	2,13	0,65
35-64 years	48%	46%	4,80	1,46
65 years or older	25%	29%	20,74	0,98
		Chi-square	27,67	> 5,991
		dof	2	
Education				
Low	29%	27%	6,08	0,80
Medium	37%	37%	0,10	1,42
High	35%	36%	1,93	0,93
		Chi-square	8,11	> 5,991
		dof	2	

Finally, after the second iteration of factor weighting, the proportions of respondents in the dataset (specifically in age, gender, and education level) were statistically indistinguishable from the population, as depicted in Table B.4. This second iteration yielded the final sample characteristics as presented in Table B.7. The factor-weighted dataset, in total of 4.416 responses, was then used for further statistical analysis.

Table B.4: Final Chi-Square Test

For 3rd Iteration				
	Expected Proportion %	Observed Proportion %	Chi-square	Weighting factor
Gender				
Man	51%	51%	0,00	1,00
Woman	49%	49%	0,00	1,00
		Chi-square	-	< 3,841
		dof	1	
Age				
18-34 years	27%	27%	0,24	0,65
35-64 years	48%	49%	0,30	1,46
65 years or older	25%	25%	0,02	0,98
		Chi-square	0,55	< 5,991
		dof	2	
Education				
Low	29%	27%	4,47	0,80
Medium	37%	37%	0,04	1,42
High	35%	36%	0,93	0,93
		Chi-square	5,44	< 5,991
		dof	2	

Table B.5: Lelylijn Project Sample Characteristics: After Factor-Weighting

Respondents' Attribute	Total Respondents*	Respondents %	Population Netherlands %
Gender			
Man	2.332	51%	51%
Woman	2.240	49%	49%
Age			
18-34 years	1.215	27%	27%
35-64 years	2.220	49%	48%
65 years or older	1.137	25%	25%
Education			
Low	1.245	27%	29%
Medium	1.684	37%	37%
High	1.643	36%	35%
Occupation			
Not working	1.773	39%	
Part-time	1.022	22%	
Full-time	1.778	39%	
Financial			
Not enough	263	6%	
Enough	3.191	70%	
More than enough	1.118	24%	
Residency			
Tenant	975	21%	
Lodger	395	9%	
Homeowner	3.202	70%	
Case specific			
Proximity			
Not proximate	2.312	51%	
Proximate	2.261	49%	
Frequency of Using Train			
Rarely	2.869	63%	
Occassionally	925	20%	
(Almost) daily	778	17%	
Frequency of Using Lelylijn			
Rarely	2.650	58%	
Occassionally	1.260	28%	
(Almost) daily	663	15%	

B.2 Oude Lijn Project

B.2.1 Sample Characteristics before Factor Weighting: Oude Lijn

After cleaning the dataset, the analysis was conducted on a total of 3.844 respondents from the Oude Lijn consultation, as presented in Table B.6. It can be observed that the composition of Oude Lijn consultation respondents differs from the general population of the Netherlands in terms of gender, age, and education. Similar to the Lelylijn analysis, a comprehensive series of statistical analyses was performed on this raw dataset. The initial results using the non-weighted dataset yielded less meaningful insights. Therefore, factor weighting was applied to the Oude Lijn dataset for further statistical analysis. The application of factor weighting resulted in slight changes in the respondents' proportions, as illustrated in Table B.7.

Table B.6: Oude Lijn Project Sample Characteristics: Before Factor-Weighting

Respondents' Attribute	Total Respondents*	Respondents %	Population Netherlands %
Gender			
Man	1.964	51%	51%
Woman	1.880	49%	49%
Age			
18-34 years	1.284	33%	27%
35-64 years	1.949	51%	48%
65 years or older	610	16%	25%
Education			
Low	521	14%	29%
Medium	1.169	30%	37%
High	2.141	56%	35%
Occupation			
Not working	1.183	31%	
Part-time	853	22%	
Full-time	1.808	47%	
Residency			
Rental private	469	12%	
Rental social	808	21%	
Owner occupied	2.567	67%	
Case specific			
Proximity			
Rest of Netherlands	1.577	41%	
South Holland	2.267	59%	
Frequency of Using Train			
Less often or never	1.753	46%	
At least a few times a month	2.091	54%	
Frequency of Using Oude Lijn			
Less often or never	2.266	59%	
At least a few times a month	1.578	41%	
Live or Work Near Station*			
Never Live/work at any station	2.311	60%	
Live/work at any station	1.533	40%	
Get in/out at Station*			
Never get in/out at any	2.307	60%	
Get in/out at any	1.537	40%	
Area of Living			
Urban	2.830	74%	
Rural	1.014	26%	

**Any station between Leiden Central, The Hague Laan of NOI, Schiedam Centrum, Dordrecht*

B.2.2 Sample Characteristics after Factor Weighting: Oude Lijn

Table B.7: Oude Lijn Project Sample Characteristics: After Factor-Weighting

Respondents' Attribute	Total Respondents*	Respondents %	Population Netherlands %
Gender			
Man	2.049	50%	51%
Woman	2.021	50%	49%
Age			
18-34 years	1.118	27%	27%
35-64 years	1.964	48%	48%
65 years or older	987	25%	25%
Education			
Low	1.115	27%	29%
Medium	1.508	37%	37%
High	1.447	36%	35%
Occupation			
Not working	1.638	40%	
Part-time	868	21%	
Full-time	1.564	38%	
Residency			
Rental private	446	11%	
Rental social	1.068	26%	
Owner occupied	2.555	63%	
Case specific			
Proximity			
Rest of Netherlands	1.915	47%	
South Holland	2.154	53%	
Frequency of Using Train			
Less often or never	2.240	55%	
At least a few times a month	1.829	45%	
Frequency of Using Oude Lijn			
Less often or never	2.748	68%	
At least a few times a month	1.322	32%	
Live or Work Near Station*			
Never Live/work at any station	2.713	67%	
Live/work at any station	1.356	33%	
Get in/out at Station*			
Never get in/out at any	2.741	67%	
Get in/out at any	1.328	33%	
Area of Living			
Urban	2.810	69%	
Rural	1.259	31%	

**Any station between Leiden Central, The Hague Laan of NOI, Schiedam Centrum, Dordrecht*

B.3 Mobility Vision Project

B.3.1 Sample Characteristics before Factor Weighting: Mobility Vision

After cleaning the dataset, the analysis was conducted on a total of 4.146 respondents from the Mobility Vision consultation, as presented in Table B.8. It can be observed that the composition of Mobility Vision consultation respondents differs from the general population of the Netherlands in terms of gender, age, and education. Similar to the Lelylijn and Oude Lijn analysis, a comprehensive series of statistical analyses was performed on this raw dataset. The initial results using the non-weighted dataset yielded less meaningful insights. Therefore, factor weighting was applied to the Mobility Vision dataset for further statistical analysis. The application of factor weighting resulted in slight changes in the respondents' proportions, as illustrated in Table B.9.

Table B.8: Mobility Vision Project Sample Characteristics: Before Factor-Weighting

Respondents' Attribute	Total Respondents*	Respondents %	Population Netherlands %
Gender			
Man	2.449	59%	51%
Woman	1.697	41%	49%
Age			
18-34 years	850	21%	27%
35-64 years	2.183	53%	48%
65 years or older	1.113	27%	25%
Education			
Low	443	11%	29%
Medium	1.227	30%	37%
High	2.476	60%	35%
Occupation			
Not working	1.462	35%	
Full-time	2.684	65%	
Area of Living			
Urban	2.588	62%	
Rural	1.558	38%	
Frequency of Facing Acessibility Problem			
Low	2.752	66%	
Moderate	1.055	25%	
High	339	8%	

B.3.2 Sample Characteristics after Factor Weighting: Mobility Vision

Table B.9: Mobility Vision Project Sample Characteristics: After Factor-Weighting

Respondents' Attribute	Total Respondents*	Respondents %	Population Netherlands %
Gender			
Man	2.161	51%	51%
Woman	2.097	49%	49%
Age			
18-34 years	1.134	27%	27%
35-64 years	2.041	48%	48%
65 years or older	1.083	25%	25%
Education			
Low	1.190	28%	29%
Medium	1.565	37%	37%
High	1.502	35%	35%
Occupation			
Not working	1.671	39%	
Full-time	2.587	61%	
Area of Living			
Urban	2.610	61%	
Rural	1.647	39%	
Frequency of Facing Acessibility Problem			
Low	2.752	65%	
Moderate	1.055	25%	
High	339	8%	

C Complete Descriptive Statistics Results

C.1 Lelylijn Project

Table C.1: Lelylijn Consultation Face Validity Evaluation

No	Face Validity Dimension	Statement	Frequency Total Sample					Mean	Std. Dev
			Totally Disagree	Disagree	Neutral	Agree	Totally Agree		
1	Transparency	The consultation steered my choices in a certain direction	20,6%	26,2%	37,4%	11,9%	4,0%	2,52	1,07
2		I trust this is a fair investigation	2,0%	3,8%	13,9%	49,2%	31,1%	4,04	0,88
3	Relevance	I thought it was an important topic to give my opinion on	0,7%	0,8%	5,5%	39,0%	54,0%	4,45	0,70
4		By participating in this consultation, I learn about the choices the government has to make about the Lely Line	2,2%	7,0%	20,5%	45,9%	24,4%	3,83	0,95
5	Acceptance	In the Netherlands we should use this method more often to involve residents in government policy	0,9%	2,3%	11,0%	38,0%	47,8%	4,29	0,82
6		If the government involves residents in choices on a large scale through this consultation, then the final decisions about the Lely Line will be more acceptable to me	3,3%	8,6%	22,4%	41,5%	24,1%	3,75	1,02
7		If the government allows residents to think about these types of choices more often, I will have more confidence in the government's decisions	3,0%	5,7%	26,2%	41,0%	24,2%	3,78	0,98

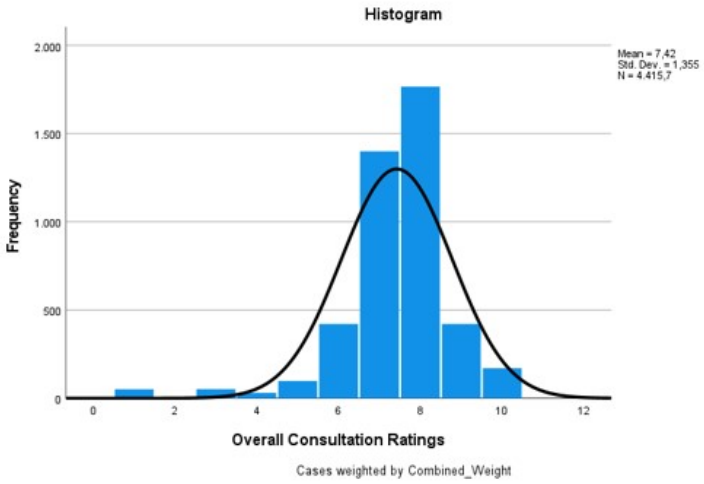


Figure C.1: Distribution of the Lelylijn Consultation's Overall Ratings

Overall Consultation Ratings		
Rating	%	Cum %
1	1,2%	1,2%
2	0,2%	1,4%
3	1,2%	2,6%
4	0,7%	3,3%
5	2,2%	5,5%
6	9,5%	15,0%
7	31,7%	46,7%
8	40,0%	86,7%
9	9,5%	96,2%
10	3,8%	100,0%
Total	**100,0%**	

Table C.2: Percentage of the Lelylijn Consultation's per Rating

C.2 Oude Lijn Project

Table C.3: Oude Lijn Consultation Face Validity Evaluation

No	Face Validity Dimension	Statement	Frequency Total Sample					Mean	Std. Dev
			Totally Disagree	Disagree	Neutral	Agree	Totally Agree		
1	Transparency	I trust this is an honest investigation	0,6%	1,4%	11,4%	47,1%	39,4%	4,23	0,75
2	Feasibility	I found the consultation difficult to understand	37,3%	34,9%	18,1%	7,0%	2,6%	2,03	1,03
3	Relevance	I thought it was an important topic to give my opinion	1,7%	3,8%	21,8%	41,9%	30,9%	3,97	0,91
4		By participating in this consultation, I have learned about the choices that the government has to make on this subject	1,5%	3,7%	24,5%	47,9%	22,4%	3,86	0,86
5	Acceptance	This method should be used more often to involve residents in public policy	0,7%	2,6%	12,6%	42,8%	41,2%	4,21	0,81
6		If many people participate in this consultation, then the final decisions on this subject are better accepted for me	1,3%	5,0%	25,8%	46,5%	21,4%	3,82	0,87
7		If the government often allows residents to think along about these types of choices more often, then I get more confidence in the decisions of the government	1,9%	4,8%	22,0%	45,6%	25,6%	3,88	0,91

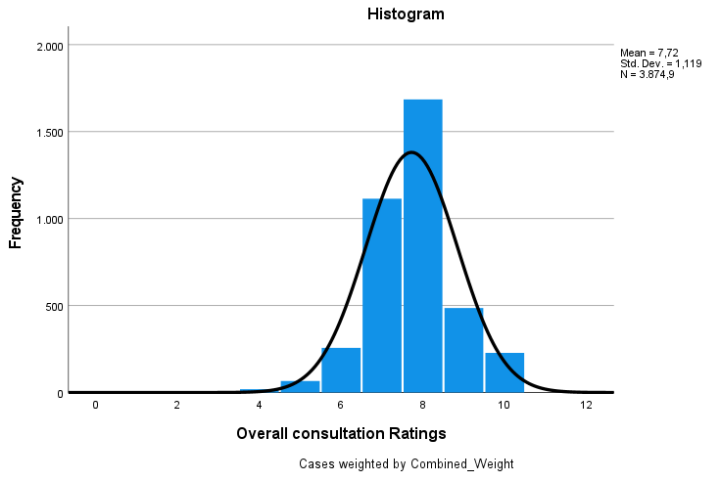


Figure C.2: Distribution of the Oude Lijn Consultation's Overall Ratings

Overall Consultation Ratings		
Rating	%	Cum %
1	0,2%	0,2%
2	0,1%	0,3%
3	0,2%	0,5%
4	0,5%	1,0%
5	1,7%	2,7%
6	6,6%	9,3%
7	28,8%	38,1%
8	43,5%	81,6%
9	12,5%	94,1%
10	5,9%	100,0%
Total	**100,0%**	

Table C.4: Percentage of the Oude Lijn Consultation's per Rating

C.3 Mobility Vision Project

Table C.5: Mobility Vision Project Consultation's Face Validity Evaluation

No	Face Validity Dimension	Statement	Frequency Total Sample					Mean	Std. Dev
			Totally	Disagree	Neutral	Agree	Totally		
1	Transparency	I trust that this research project is conducted in a fair way	0,8%	1,8%	12,6%	47,7%	37,2%	4,19	0,78
2	Feasibility	I found the consultation difficult to understand	37,2%	35,2%	17,7%	7,4%	2,4%	2,03	1,03
3	Relevance	I thought it was an important topic to give my opinion on	0,5%	2,0%	12,5%	47,9%	37,0%	4,19	0,77
4	Relevance	By participating in this consultation, I learned about the choices the Province must make on this topic	2,7%	6,3%	29,6%	45,8%	15,6%	3,65	0,91
5	Acceptance	This method should be used more often to involve residents in government policies	1,0%	2,0%	17,7%	45,7%	33,7%	4,09	0,82
6	Acceptance	If many people participate in this consultation, the final decisions on this topic are more acceptable to me	2,1%	6,2%	30,7%	44,6%	16,4%	3,67	0,89
7	Acceptance	If the government involves residents more often in thinking about these kinds of choices, I will have more confidence in the government's decisions	2,0%	5,4%	26,8%	45,5%	20,2%	3,77	0,90

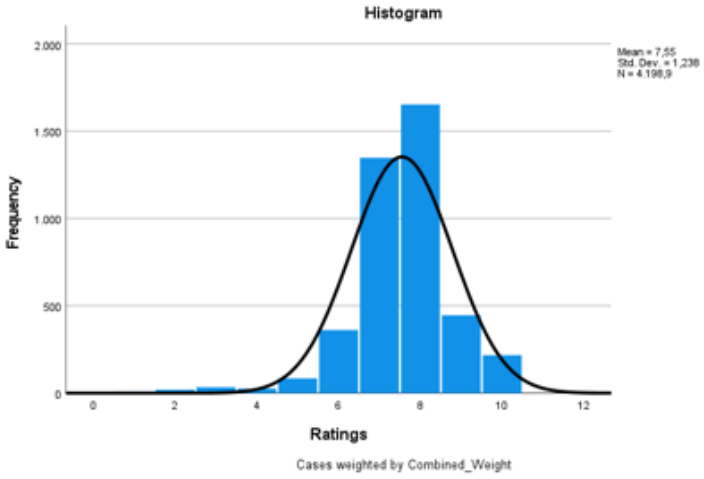


Figure C.3: Distribution of the Mobility Vision Consultation's Overall Ratings

Overall Consultation Ratings		
Rating	%	Cum %
1	0,2%	0,2%
2	0,5%	0,7%
3	0,8%	1,5%
4	0,6%	2,1%
5	2,0%	4,1%
6	8,6%	12,7%
7	32,1%	44,8%
8	39,4%	84,2%
9	10,6%	94,8%
10	5,2%	100,0%
Total	**100%**	

Table C.6: Percentage of the Mobility Vision Consultation's per Rating

D Complete Latent Class Cluster Analysis (LCCA) Results

D.1 Lelylijn Project

Table D.1: Lelylijn Consultation LCCA Results: Indicators

	Cluster1	Cluster2	Cluster3	Cluster4	Wald	p-value
Cluster Size	54,3%	24,4%	11,1%	10,3%		
Indicators						
Statement 1: The consultation steered my choices in a certain direction						
Disagree	55,0%	49,4%	9,5%	28,0%		
Neutral	36,4%	39,4%	37,6%	45,6%	297,29	3,8E-64
Agree	8,6%	11,2%	52,9%	26,4%		
Statement 2: I trust this is a fair investigation						
Disagree	0,3%	0,8%	14,6%	35,9%		
Neutral	8,1%	12,3%	39,0%	42,1%	571,61	1,4E-123
Agree	91,6%	87,0%	46,4%	22,0%		
Statement 3: I thought it was an important topic to give my opinion on						
Disagree	0,1%	1,5%	0,2%	10,4%		
Neutral	2,6%	9,2%	3,8%	21,0%	210,67	2,1E-45
Agree	97,3%	89,3%	96,0%	68,6%		
Statement 4: In the Netherlands we should use this method more often to involve residents in government policy						
Disagree	0,0%	4,5%	1,0%	18,5%		
Neutral	1,8%	22,7%	11,7%	36,1%	317,04	2,0E-68
Agree	98,2%	72,8%	87,3%	45,4%		
Statement 5: By participating in this consultation, I learn about the choices the government has to make about the Lelylijn						
Disagree	0,7%	7,6%	20,7%	42,7%		
Neutral	11,5%	31,6%	41,4%	40,1%	623,49	8,1E-135
Agree	87,8%	60,8%	37,9%	17,3%		
Statement 6: If the government involves residents in choices on a large scale through this consultation, then the final decisions about the Lelylijn will be more acceptable to me						
Disagree	0,1%	19,5%	4,6%	64,6%		
Neutral	6,6%	50,5%	34,8%	31,8%	444,19	5,9E-96
Agree	93,3%	30,0%	60,6%	3,6%		
Statement 7: If the government allows residents to think about these types of choices more often, I will have more confidence in the government's decisions						
Disagree	0,2%	13,8%	6,0%	46,9%		
Neutral	9,2%	52,8%	43,1%	45,7%	474,10	1,9E-102
Agree	90,6%	33,5%	50,9%	7,4%		
Overall Ratings						
1	0,0%	0,0%	2,0%	6,3%		
2	0,0%	0,0%	0,6%	1,4%		
3	0,0%	0,0%	2,7%	5,3%		
4	0,0%	0,0%	2,8%	4,4%		
5	0,0%	0,1%	11,1%	13,7%		
6	1,4%	4,7%	39,6%	38,2%	405,25	1,6E-87
7	28,2%	46,4%	37,9%	28,8%		
8	51,8%	42,3%	3,4%	2,0%		
9	13,2%	5,3%	0,0%	0,0%		
10	5,4%	1,1%	0,0%	0,0%		
Mean	7,93	7,51	6,08	5,58		

Table D.2: Lelylijn Consultation LCCA Results: Covariates

	Cluster1	Cluster2	Cluster3	Cluster4	Wald	p-value
Cluster Size	54,3%	24,4%	11,1%	10,3%		
Covariates						
Age						
18-34 years	29,8%	25,4%	15,8%	24,1%		
35-64 years	48,4%	44,6%	47,9%	59,6%	47,74	2,4E-10
65 years or older	21,8%	30,1%	36,3%	16,3%		
Gender						
Man	51,4%	48,1%	53,1%	53,5%	1,36	7,2E-01
Woman	48,6%	51,9%	46,9%	46,5%		
Education						
Low	28,2%	22,0%	31,5%	29,9%		
Medium	40,1%	33,4%	35,4%	29,5%	44,76	1,0E-09
High	31,7%	44,6%	33,1%	40,7%		
Occupation						
Not working	36,4%	44,1%	46,5%	30,2%		
Part-time	23,9%	20,4%	14,6%	27,2%	15,79	1,3E-03
Full-time	39,7%	35,5%	39,0%	42,6%		
Proximity						
Not proximate	48,5%	46,3%	69,0%	51,8%	54,49	8,8E-12
Proximate	51,5%	53,7%	31,0%	48,3%		
Train Use Frequency						
Rarely	60,2%	63,8%	60,7%	76,0%		
Occassionally	22,5%	17,1%	21,5%	14,6%	23,20	3,7E-05
(Almost) daily	17,3%	19,1%	17,8%	9,4%		
Lelylijn (plan to) Use Frequency						
Rarely	55,5%	61,9%	50,9%	69,2%		
Occassionally	27,7%	26,9%	32,6%	22,8%	38,83	1,9E-08
(Almost) daily	16,8%	11,2%	16,5%	8,1%		
Residence						
Tenant	22,5%	23,7%	14,2%	17,3%		
Lodger	10,5%	6,5%	3,3%	10,0%	26,35	8,0E-06
Homeowner	67,1%	69,8%	82,6%	72,7%		
Financial Level						
Not enough	5,2%	6,5%	7,6%	4,9%		
Enough	70,7%	64,4%	78,3%	68,8%	30,60	1,0E-06
More than enough	24,1%	29,2%	14,2%	26,3%		

D.2 Oude Lijn Project

Table D.3: Oude Lijn Consultation LCCA Results: Indicators

	Cluster1	Cluster2	Cluster3	Cluster4	Wald	p-value
Cluster Size	48,4%	20,7%	18,1%	12,8%		
Indicators						
Statement 1: I trust this is an honest investigation						
Disagree	0,0%	0,4%	1,0%	14,1%		
Neutral	2,1%	10,9%	16,2%	43,8%	367,47	2,5E-79
Agree	97,9%	88,7%	82,8%	42,1%		
Statement 2: I thought it was an important topic to give my opinion						
Disagree	0,0%	12,1%	0,1%	20,9%		
Neutral	4,2%	55,2%	5,8%	58,1%	249,81	7,2E-54
Agree	95,8%	32,6%	94,2%	21,0%		
Statement 3: I found the consultation difficult to understand						
Disagree	76,6%	71,2%	84,7%	48,3%		
Neutral	16,8%	19,3%	12,1%	26,1%	177,16	3,6E-38
Agree	6,6%	9,5%	3,1%	25,6%		
Statement 4: This method should be used more often to involve residents in public policy						
Disagree	0,0%	0,2%	1,0%	19,0%		
Neutral	1,5%	9,6%	20,1%	53,2%	411,58	6,8E-89
Agree	98,6%	90,2%	79,0%	27,8%		
Statement 5: By participating in this consultation, I have learned about the choices that the government has to make on this subject						
Disagree	0,1%	2,0%	9,1%	23,9%		
Neutral	7,2%	28,9%	48,2%	54,9%	429,99	7,0E-93
Agree	92,7%	69,1%	42,7%	21,2%		
Statement 6: If many people participate in this consultation, then the final decisions on this subject are better accepted for me						
Disagree	0,1%	1,8%	11,4%	27,7%		
Neutral	6,2%	29,7%	53,8%	56,5%	529,79	1,7E-114
Agree	93,8%	68,6%	34,8%	15,9%		
Statement 7 If the government often allows residents to think along about these types of choices more often, then I get more confidence in the decisions of the government						
Disagree	0,0%	0,8%	11,7%	30,5%		
Neutral	3,7%	20,0%	52,1%	54,4%	421,27	5,5E-91
Agree	96,3%	79,2%	36,2%	15,2%		
Ratings						
1	0,0%	0,0%	0,0%	1,9%		
2	0,0%	0,0%	0,0%	0,9%		
3	0,0%	0,1%	0,1%	1,9%		
4	0,0%	0,2%	0,2%	2,9%		
5	0,1%	1,5%	1,4%	8,8%		
6	1,5%	8,6%	8,1%	21,3%	472,20	5,0E-102
7	17,7%	38,8%	38,1%	41,3%		
8	50,9%	43,1%	43,9%	19,7%		
9	19,6%	6,4%	6,8%	1,3%		
10	10,2%	1,3%	1,4%	0,1%		
Mean	8,19	7,47	7,50	6,52		

Table D.4: Oude Lijn Consultation LCCA Results: Covariates

	Cluster1	Cluster2	Cluster3	Cluster4	Wald	p-value
Cluster Size	48,4%	20,7%	18,1%	12,8%		
Covariates						
Age						
18-34 years	28,9%	25,7%	26,5%	26,1%		
35-64 years	47,4%	47,2%	50,5%	50,7%	18,00	0,00
65 years or older	23,7%	27,1%	23,0%	23,3%		
Gender						
Man	53,5%	39,2%	54,9%	49,8%	9,62	0,02
Woman	46,5%	60,8%	45,1%	50,2%		
Education						
Low	26,9%	35,0%	11,7%	40,2%		
Medium	37,5%	44,9%	29,1%	35,3%	34,42	0,00
High	35,7%	20,2%	59,2%	24,5%		
Occupation						
Not working	39,0%	46,5%	31,5%	47,2%		
Part-time	19,8%	26,4%	19,9%	20,9%	5,00	0,17
Full-time	41,3%	27,1%	48,7%	31,9%		
Proximity						
Rest of the Netherlands	44,0%	65,8%	21,7%	64,1%	55,13	0,00
South Holland	56,0%	34,2%	78,3%	35,9%		
Train Use Frequency						
At least a few times a month	51,7%	21,9%	63,6%	30,7%	28,44	0,00
Less often or never	48,3%	78,1%	36,4%	69,4%		
Oude Lijn (plan to) Use Frequency						
At least a few times a month	38,6%	7,2%	52,0%	22,4%	15,44	0,00
Less often or never	61,4%	92,8%	48,0%	77,6%		
In/out from any Station*						
In/ out at any station	37,8%	10,3%	51,5%	22,5%	1,32	0,72
Never in/out at any station	62,2%	89,7%	48,6%	77,5%		
Live/work near any Station*						
Live/work at any station	37,9%	11,8%	49,5%	27,4%	14,26	0,00
Never Live/work at any station	62,1%	88,2%	50,5%	72,6%		
Area of Live						
Urban	72,6%	54,8%	84,3%	57,1%	21,47	0,00
Rural	27,4%	45,2%	15,7%	42,9%		
Type of Home						
Rental private	11,7%	9,8%	9,7%	12,2%		
Rental social	26,6%	28,0%	17,2%	34,8%	10,58	0,01
Owner occupied	61,7%	62,3%	73,2%	53,0%		

*Any station between Leiden Central, The Hague Laan of NOI, Schiedam Centrum, Dordrecht

D.3 Mobility Vision Project

Table D.5: Mobility Vision Consultation LCCA Results: Indicators

	Cluster1	Cluster2	Cluster3	Cluster4	Wald	p-value
Cluster Size	50,5%	30,0%	14,6%	4,9%		
Indicators						
Statement 1: I trust that this research project is conducted in a fair way						
Disagree	0,0%	0,4%	8,2%	19,0%		
Neutral	3,0%	10,7%	39,5%	47,8%	454,30	4E-98
Agree	97,0%	88,9%	52,3%	33,1%		
Statement 2: I thought it was an important topic to give my opinion						
Disagree	0,0%	0,2%	12,0%	9,7%		
Neutral	2,1%	8,2%	43,5%	41,2%	456,04	2E-98
Agree	97,9%	91,6%	44,5%	49,2%		
Statement 3: I found the consultation difficult to understand						
Disagree	76,1%	88,0%	43,2%	67,1%		
Neutral	17,6%	10,2%	28,2%	21,8%	240,17	9E-52
Agree	6,3%	1,8%	28,6%	11,1%		
Statement 4: This method should be used more often to involve residents in government policies						
Disagree	0,0%	1,0%	9,8%	24,3%		
Neutral	2,5%	20,6%	47,4%	53,6%	469,21	2E-101
Agree	97,5%	78,4%	42,9%	22,1%		
Statement 5: By participating in this consultation, I learned about the choices the Province must make on this topic						
Disagree	0,4%	11,3%	13,6%	62,3%		
Neutral	13,5%	48,5%	50,3%	34,1%	444,07	6E-96
Agree	86,1%	40,2%	36,1%	3,6%		
Statement 6: If many people participate in this consultation, the final decisions on this topic are more acceptable to me						
Disagree	0,0%	9,3%	8,1%	79,0%		
Neutral	6,8%	61,0%	59,5%	20,6%	510,91	2E-110
Agree	93,2%	29,7%	32,3%	0,4%		
Statement 7: If the government involves residents more often in thinking about these kinds of choices, I will have more confidence in the government's decisions						
Disagree	0,0%	8,3%	8,6%	75,2%		
Neutral	2,9%	54,3%	54,7%	24,0%	380,88	3E-82
Agree	97,1%	37,4%	36,7%	0,8%		
Ratings						
1	0,0%	0,0%	0,8%	2,9%		
2	0,0%	0,0%	1,2%	3,4%		
3	0,0%	0,0%	2,6%	5,6%		
4	0,0%	0,1%	2,4%	4,1%		
5	0,1%	0,9%	8,4%	11,2%		
6	1,6%	8,0%	24,6%	25,9%	543,65	2E-117
7	21,8%	43,7%	45,2%	37,4%		
8	51,0%	40,7%	14,2%	9,2%		
9	17,4%	5,5%	0,7%	0,3%		
10	8,2%	1,1%	0,0%	0,0%		
Mean	8,09	7,45	6,46	5,93		

Table D.6: Mobility Vision Consultation LCCA Results: Covariates

	Cluster1	Cluster2	Cluster3	Cluster4	Wald	p-value
Cluster Size	50,5%	30,0%	14,6%	4,9%		
Covariates						
Age						
18-34 years	25,6%	21,4%	41,0%	26,5%		
35-64 years	47,7%	50,9%	40,4%	54,7%	43,52	0,00
65 years or older	26,7%	27,8%	18,6%	18,9%		
Gender						
Man	52,2%	51,2%	40,4%	63,5%	15,38	2E-03
Woman	47,8%	48,8%	59,6%	36,6%		
Education						
Low	30,0%	17,1%	46,1%	18,8%		
Medium	39,6%	33,2%	37,5%	27,0%	149,14	0,00
High	30,4%	49,7%	16,4%	54,2%		
Area_Live						
Urban	60,6%	62,2%	59,2%	69,9%	3,42	0,33
Rural	39,5%	37,8%	40,8%	30,1%		
Frequency of Facing Accessibility Problem						
Low	71,0%	68,5%	63,8%	60,0%		
Medium	21,5%	24,4%	27,5%	27,5%	7,46	0,06
High	7,5%	7,1%	8,7%	12,5%		
Occupation						
Not working	40,3%	38,1%	40,1%	32,9%	0,21	0,98
Full-time	59,7%	61,9%	59,9%	67,1%		

E Test of Parallel Lines

Parallel line tests were conducted to evaluate the proportional odds assumption before performing ordinal logistic regression. The objective of this test is to determine whether the assumption of proportional odds holds true in an ordinal logistic regression model. Specifically, in this context, the outcome groups (i.e., ”agree,” ”neutral,” and ”disagree”) are examined. This test assesses if the relationship between each pair of these outcome groups is statistically the same, indicating that the effect of the predictor variables is consistent across different threshold levels of the ordinal outcome group. The significance of the chi-square statistic in the test of parallel lines is used to check this assumption. A significant p-value (less than 0,05) suggests that the assumption of proportional odds (parallel lines) is violated. In other words, the relationship between the predictors and the outcome variable differs across the response categories of ”agree,” ”neutral,” and ”disagree.” This violation indicates that the proportional odds assumption may not be appropriate for the data, and hence the use of ordinal logistic regression might not be suitable. Therefore, if the p-value is significant, the use of multinomial logistic regression, which does not assume proportional odds, might be more appropriate.

E.1 Lelylijn Project

Table E.1 presents the results of the parallel line tests for all statements in the Oude Lijn dataset.

Table E.1: Test of Parallel Lines of Lelylijn Consultation Dataset

No	Statement	STEP 1		STEP 2	
		Chi-square	Sig.	Chi-square	Sig.
1	The consultation steered my choices in a certain direction	150,697	<,001	222,069	<,001
2	I trust this is a fair investigation	47,176*	<,001	56,593*	<,001
3	I thought it was an important topic to give my opinion on	.*	.	.*	.
4	In the Netherlands we should use this method more often to involve residents in government policy	37,586*	<,001	11,615*	0,77
5	By participating in this consultation, I learn about the choices the government has to make about the Lely Line	40,314*	<,001	56,970*	<,001
6	If the government involves residents in choices on a large scale through this consultation, then the final decisions about the Lely Line will be more acceptable to me	92,704*	<,001	127,712*	<,001
7	If the government allows residents to think about these types of choices more often, I will have more confidence in the government's decisions	29,296*	0,004	83,436*	<,001

*The Chi-Square statistic is computed based on the log-likelihood value of the last iteration of the general model. Validity of the test is uncertain.

E.2 Oude Lijn Project

Table E.2 presents the results of the parallel line tests for all statements in the Oude Lijn dataset. The findings indicate that the test results for all statements in step 1 (i.e., using only respondents' main characteristics) and for almost all statements in step 2 (i.e., using all respondents' characteristics as independent variables) are statistically significant. These results suggest that the assumption of proportional odds is violated, implying that ordinal logistic regression may not be appropriate for further analysis. Consequently, a multinomial logistic regression was conducted for each statement.

Table E.2: Test of Parallel Lines of Oude Lijn Consultation Dataset

No	Statement	STEP 1		STEP 2	
		Chi-square	Sig.	Chi-square	Sig.
1	I trust this is an honest investigation	18,585*	0,01	25,787*	0,04
2	I found the consultation difficult to understand	177,208	<,001	329,924	<,001
3	I thought it was an important topic to give my opinion	31,344	<,001	19,545*	0,19
4	By participating in this consultation, I have learned about the choices that the government has to make on this subject	37,571	<,001	52,171*	<,001
5	This method should be used more often to involve residents in public policy	39,95	<,001	56,622*	<,001
6	If many people participate in this consultation, then the final decisions on this subject are better accepted for me	31,764	<,001	45,574	<,001
7	If the government often allows residents to think along about these types of choices more often, then I get more confidence in the decisions of the government	24,919	<,001	31,869*	0,01

*The Chi-Square statistic is computed based on the log-likelihood value of the last iteration of the general model. Validity of the test is uncertain.

E.3 Mobility Vision Project

Table E.3 presents the results of the parallel line tests for all statements in the Mobility Vision dataset. The findings indicate that the test results for almost all statements in step 1 (i.e., using only respondents' main characteristics) and in step 2 (i.e., using all respondents' characteristics as independent variables) are statistically significant. These results suggest that the assumption of proportional odds is violated, implying that ordinal logistic regression may not be appropriate for further analysis. Consequently, a multinomial logistic regression was conducted for each statement.

Table E.3: Test of Parallel Lines of Mobility Vision Consultation Dataset

No	Statement	STEP 1		STEP 2	
		Chi-square	Sig.	Chi-square	Sig.
1	The consultation steered my choices in a certain direction	31,397	<,001	9,133*	0,425
2	I trust this is a fair investigation	.*	.	.*	.
3	I thought it was an important topic to give my opinion on	333,615	<,001	347,184	<,001
4	In the Netherlands we should use this method more often to involve residents in government policy	40,317*	<,001	14,888*	0,094
5	By participating in this consultation, I learn about the choices the government has to make about the Lely Line	82,368	<,001	82,709	<,001
6	If the government involves residents in choices on a large scale through this consultation, then the final decisions about the Lely Line will be more acceptable to me	96,509	<,001	87,215*	<,001
7	If the government allows residents to think about these types of choices more often, I will have more confidence in the government's decisions	38,782	<,001	36,464*	<,001

*The Chi-Square statistic is computed based on the log-likelihood value of the last iteration of the general model. Validity of the test is uncertain.

F Coding for Logistic Regression

This chapter presents the coding for respondents' characteristics or answers used in the analysis. All "I'd rather not say" or "I don't know" responses have been removed to streamline the analysis process, considering that the dataset remains sufficiently large even after their removal.

F.1 Lelylijn

Table F.1: Coding for Multinomial Logistic Regression: Lelylijn part 1

General Demographics			
Variable	**Actual options in the consultation**	**Categories**	**Coding**
Gender	I am a woman	Woman	2
	I'd rather not say / I don't know	I'd rather not say	*deleted*
	I'm a man	Man	1
Age	Younger than 25 years old	18-34 years	1
	Between 25 and 34 years	18-34 years	1
	Between 35 and 44 years	35-64 years	2
	Between 45 and 54 years	35-64 years	2
	Between 55 and 64 years	35-64 years	2
	65 years or older	65 years or older	3
	I'd rather not say / I don't know	I'd rather not say	*deleted*
Education	Class 1, 2 or 3 HAVO/VWO	Low	1
	Class 4, 5 or 6 HAVO/VWO	Medium	2
	HBO	High	3
	MBO level 2, 3 or 4 (basic vocational training, vocational training, or middle management and specialist training)	Medium	2
	Primary school	Low	1
	University	High	3
	VMBO, MBO level 1	Low	1
	I'd rather not say / I don't know	I'd rather not say	*deleted*
Occupation	Combination of study and part-time job	Part-time	2
	Entrepreneur (with employees)	Full-time	3
	I'd rather not say that	I'd rather not say	*deleted*
	Early retirement/retired	Not working	1
	Entrepreneur (self-employed)	Full-time	3
	Full-time employed (≥35 hours)	Full-time	3
	HBO/WO student	Not working	1
	High school student	Not working	1
	Housewife or husbandhusband	Not working	1
	I'd rather not say / I don't know	I'd rather not say	*deleted*
	Looking for work	Not working	1
	MBO student	Not working	1
	Otherwise	I'd rather not say	1
	Part-time employed (< 35 hours)	Part-time	2
	Volunteer	I'd rather not say	*deleted*
Financial	have enough money every month	Enough money	2
	I don't have enough money every month	Not enough money	1
	I have enough money every month	Enough money	2
	I have more than enough money every month	More than enough money	3
	I'd rather not say / I don't know	I'd rather not say	*deleted*

Table F.2: Coding for Multinomial Logistic Regression: Lelylijn part 2

General Demographics			
Variable	**Actual options in the consultation**	**Categories**	**Coding**
Residency	I live in a rental property	Tenant	1
	I live in an owner-occupied home	Homeowner	3
	I live in someone's house	Lodger	2
	I'd rather not say / I don't know	Lodger	2
	live in an owner-occupied home	Homeowner	3
Proximity	Drenthe	Not proximate	1
	Flevoland	Proximate	2
	Friesland	Not proximate	1
	Gelderland	Not proximate	1
	Groningen	Proximate	2
	I'd rather not say	Not proximate	1
	Limburg	Not proximate	1
	North Brabant	Not proximate	1
	North Holland	Proximate	2
	Overijssel	Not proximate	1
	South-Holland	Not proximate	1
	Utrecht	Not proximate	1
	Zealand	Not proximate	1
Case Specific Characteristics			
Frequency to Use Train/ Lelylijn	-	I'd rather not say	*deleted*
	2-4 days a week	(Almost) Daily	3
	5-7 days a week	(Almost) Daily	3
	Daily or almost daily	(Almost) Daily	3
	Never or almost never	Rarely	1
	Several times a month	Occassionally	2
	Several times a week	(Almost) Daily	3
	Several times a year	Rarely	1
Face Validity Evaluation			
Face Validity	Totally disagree	Disagree	1
	Disagree	Disagree	2
	Neutral	Neutral	3
	Agree	Agree	4
	Totally agree	Agree	5
	I'd rather not say/I don't know	I'd rather not say	*deleted*

F.2 Oude Lijn

Table F.3: Coding for Multinomial Logistic Regression: Oude Lijn

General Demographics		
Variable	**Categories**	**Coding**
Age	Younger than 35	[lft3=1]
	Between 35 and 65	[lft3=2]
	Older than 65	[lft3=3]
Gender	Man	[gender=1]
	Woman	[gender=2]
Education	Low	[opl3=1]
	Medium	[opl3=2]
	High	[opl3=3]
Occupation	Not working	[occupation_coded=1]
	Part-time	[occupation_coded=2]
	Full-time	[occupation_coded=3]
Case-specific Characteristics		
Variable	**Categories**	**Coding**
Province	Rest of Netherlands	[dzh=0]
	South Holland	[dzh=1]
Frequency of Train Use	At least a few times a month	[trein=1]
	Less often or never	[trein=2]
Frequency of Oudelijn Use	At least a few times a month	[ol=1]
	Less often or never	[ol=2]
Get in/out at any of 4 Stations	In/ out at any station	[oliu=1]
	Never in/out at any station	[oliu=2]
Live/ work near any of 4 Stations	Live/work at any station	[olww=1]
	Never Live/work at any station	[olww=2]
Area of Living	Urban	[live_area=1]
	Rural	[live_area=2]
Residency	Rental property in the private sector	[type_home=1]
	Rental property in the social rental sector	[type_home=2]
	Owner-occupied home	[type_home=3]

F.3 Mobility Vision

Table F.4: Coding for Multinomial Logistic Regression: Mobility Vision

General Demographics			
Variable	**Actual options in the consultation**	**Categories**	**Coding**
Age	35-44	Between 35 and 65	2
	45-54	Between 35 and 65	2
	55-64	Between 35 and 65	2
	65 or older	Older than 65	3
	17 years or younger	Under 35	1
	18-24	Under 35	1
	25-34	Under 35	1
Area of Live	A rural area	Rural	2
	A small community	Rural	2
	A big city	Urban	1
	A little city	Urban	1
Education	HBO	High	3
	University	High	3
	HAVO/VWO class 1, 2 or 3	Low	1
	MBO level 1	Low	1
	Primary school	Low	1
	Vmbo	Low	1
	HAVO/VWO class 4, 5 or 6	Middle	2
	MBO level 2, 3 or 4 (basic vocational training, vocational training, or middle management and specialist training)	Middle	2
Gender	Man	Man	1
	Woman	Woman	2
Occupation	I do paid work	Full-time	2
	I am a househusband/housewife	Not working	1
	I am retired	Not working	1
	I do volunteer work	Not working	1
	I go to school/study	Not working	1
Case-specific Characteristics			
Frequency of Problems with Accessibility	Often (a few times a week)	High	3
	Very often (every day)	High	3
	Occasionally (a few times a year)	Low	1
	Regularly (a few times a month)	Moderate	2

G Complete Results of Multinomial Logistic Regression

This chapter presents the comprehensive outcomes of the Multinomial Logistic Regressions and analysis conducted for the Oude Lijn and New Mobility Project consultation. Furthermore, a summary table displaying statistical findings for the Non-weighted Model of the Lelylijn dataset is provided as supplementary material to explain the rationale behind selecting the weighted model over the non-weighted model.

G.1 Statistics Summary of Non-weighted Model Results

G.1.1 Lelylijn

Table G.1: Statistics Results Summary of multinomial logistic regression for all face validity statement using non-weighted model

Statistics	Statement						
	1	2	3	4	5	6	7
Likelihood Ratio Test	<,001	<,001	<,001	<,001	<,001	<,001	<,001
Pearson	0,548	0,006	0,486	0,498	0,155	0,365	0,313
Deviance	0,018	1	1	1	1	0,839	0,831
Pseudo R-Square: Cox and Snell	0,024	0,037	0,028	0,024	0,011	0,018	0,023
Pseudo R-Square: Nagelkerke	0,028	0,051	0,054	0,029	0,017	0,021	0,027
Pseudo R-Square: McFadden	0,012	0,028	0,039	0,014	0,01	0,01	0,013
Correct Percentage	49,80%	77,30%	90,60%	66,00%	82,60%	62,70%	60,10%

G.2 Complete Weighted Model Results: Oude Lijn

The Multinomial Logistic Regression analysis was conducted for each face validity statement in Oude Lijn consultation. The analysis was conducted with three outcome groups: (Totally) Agree, Neutral, and (Totally) Disagree. The regression was performed twice, using ”Neutral” and ”(Totally) Agree” as the reference outcomes, resulting in three comparisons: (Totally) Agree vs. Neutral, (Totally) Disagree vs. Neutral, and (Totally) Disagree vs. (Totally) Agree.

The ”B” column represents the estimated coefficients for the predictor variables. When the predictors are categorical, the coefficients indicate the change in the log-odds of the outcome categories relative to the reference category of the predictor variable. The ”Sig.” column in the results table indicates the significance level (p-value) for the test of the null hypothesis that the corresponding coefficient is equal to zero. If the p-value is less than 0.05, the null hypothesis is rejected, suggesting that the predictor variable significantly affects the outcome variable. Conversely, if the p-value is greater than or equal to 0.05, there is not enough evidence to suggest a significant effect of the predictor variable on the outcome variable.

G.2.1 I trust this is an honest investigation

The first multinomial logistic regression conducted on the Oude Lijn dataset was on the statement ”*I trust this is an honest investigation*”, as presented in Table G.2. The analysis revealed that 5 attributes were significant in one or more comparisons when controlled for other independent variables.

Table G.2: Results of multinomial logistic regression for statement: I trust this is an honest investigation

Attributes	Agree vs Neutral		Disagree vs Neutral		Disagree vs Agree	
	B	Sig.	B	Sig.	B	Sig.
Intercept	1,903	<,001	-2,934	<,001	-4,838	<,001
Age						
Younger than 35	0,142	0,393	0,665	0,152	0,524	0,235
Between 35 and 65	-0,156	0,333	0,944	0,037	1,100	0,011
Older than 65	REFERENCE					
Gender						
Man	0,094	0,391	0,531	0,047	0,437	0,081
Woman	REFERENCE					
Education Level						
Low	0,066	0,652	-0,575	0,128	-0,641	0,072
Medium	0,202	0,163	0,232	0,476	0,03	0,921
High	REFERENCE					
Occupation						
Not working	-0,048	0,734	0,328	0,323	0,376	0,221
Part-time	0,006	0,969	0,018	0,96	0,013	0,97
Full-time	REFERENCE					
Province						
Rest of Netherlands	-0,137	0,237	-0,044	0,878	0,094	0,726
South Holland	REFERENCE					
Frequency of Train Use						
At least a few times a month	0,412	0,01	-0,342	0,393	-0,754	0,045
Less often or never	REFERENCE					
Frequency of Oudelijn Use						
At least a few times a month	-0,354	0,076	-0,032	0,946	0,323	0,454
Less often or never	REFERENCE					
Get in/ out at any station*						
In/ out at any station	0,397	0,031	0,926	0,029	0,529	0,178
Never in/out at any station	REFERENCE					
Live/Work near any station*						
Live/work at any station	-0,061	0,645	-0,033	0,917	0,028	0,923
Never Live/work at any station	REFERENCE					
Area of Living						
Urban	0,129	0,267	-0,149	0,604	-0,278	0,303
Rural	REFERENCE					
Residency						
Rental property in the private sector	-0,329	0,039	0,258	0,492	0,587	0,092
Rental property in the social rental	-0,426	<,001	0,34	0,251	0,766	0,006
Owner-occupied home	REFERENCE					
Likelihood Ratio Test	<,001					
Pearson	<,001					
Deviance	1					
Pseudo R-Square: Cox and Snell	0,021					
Pseudo R-Square: Nagelkerke	0,035					
Pseudo R-Square: McFadden	0,023					
Correct Percentage	86,7%					

*Any station between Leiden Central, The Hague Laan of NOI, Schiedam Centrum, Dordrecht

Within this statement, it can be observed that:

1. **Age:**
   Middle-aged people (35-65) are more likely to (Totally) Disagree than be Neutral (coefficient = 0.944, p = 0.037) or (Totally) Disagree than (Totally) Agree (coefficient = 1.1, p = 0.011) with the statement, relative to elderly people (65+). This indicates that middle-aged people are less likely to agree or be neutral about the honesty of the investigation compared to seniors.
2. **Gender:**
   Men are more likely to (Totally) Disagree than be Neutral with this statement compared to women (coefficient = 0.531, p = 0.047).
3. **Frequency of Train Use:**
   People who travel by train at least a few times a month are more likely to (Totally) Agree than be Neutral (coefficient = 0.412, p = 0.01) and less likely to (Totally) Disagree than (Totally) Agree

(coefficient = -0.754, p = 0.045) compared to people who travel less frequently by train.

4. **Get in/out at any station:**
   People who ever get in or out at any of the four stations mentioned in the consultation (i.e., Leiden Central, The Hague Laan van NOI, Schiedam Centrum, Dordrecht) are more likely to (Totally) Agree (coefficient = 0.397, p = 0.031) and more likely to (Totally) Disagree (coefficient = 0.926, p = 0.029) than be Neutral with this statement compared to people who never get in or out at those stations. This suggests that those who use the specified stations have polarized opinions about the honesty of the investigation, being more likely to have a strong opinion (either positive or negative) rather than being neutral.
5. **Residency:**
   People who live in rental property in the private sector are less likely to (Totally) Agree than be Neutral (coefficient = -0.329, p = 0.039) compared to people who live in an owner-occupied house. People who live in rental property in the social rental sector are less likely to (Totally) Agree than be Neutral (coefficient = -0.426, p < 0.001) and more likely to (Totally) Disagree than (Totally) Agree (coefficient = 0.766, p = 0.006) compared to people who live in an owner-occupied house.

The Likelihood Ratio Test was significant ($p < 0.001$), indicating the model's superiority over the intercept-only model in predicting whether the consultation is perceived as honest research. One of either Pearson or Deviance is non-significant ($p > 0.05$), suggesting the model's good fit to the data. The Pseudo R-Square values (Cox and Snell, Nagelkerke, McFadden) suggest that the model explains a small proportion of the variance in the outcome variable. Nevertheless, the model can still correctly classifies 86.7% of cases.

In conclusion, in evaluating whether this consultation is an honest investigation, the analysis shows that middle-aged people are more likely to (Totally) Disagree with the idea compared to seniors, indicating less positive evaluation. People who travel by train occasionally are more likely to (Totally) Agree that this is an honest investigation compared to those who travel less frequently. People who ever get in or out at the studied stations are more likely to (Totally) Agree or (Totally) Disagree rather than be Neutral, suggesting polarized opinions among station users. Finally, people who live in rental properties are less likely to think that this is an honest investigation than people who live in owner-occupied houses.

G.2.2 I thought it was an important topic to give my opinion

The second multinomial logistic regression on Oude Lijn dataset was conducted on the statement "*I thought it was an important topic to give my opinion*" and presented presented in Table G.3. The analysis revealed 9 significant attributes when controlled for other independent variables.

Table G.3: Results of multinomial logistic regression for statement: I thought it was an important topic to give my opinion

Attributes	Agree vs Neutral		Disagree vs Neutral		Disagree vs Agree	
	B	Sig.	B	Sig.	B	Sig.
Age						
Younger than 35	-0,401	0,002	1,036	<,001	1,437	<,001
Between 35 and 65	-0,259	0,045	0,789	0,006	1,048	<,001
Older than 65	REFERENCE					
Gender						
Man	0,112	0,205	-0,027	0,877	-0,139	0,391
Woman	REFERENCE					
Education Level						
Low	-0,457	<,001	-0,569	0,013	-0,112	0,604
Medium	-0,069	0,561	-0,02	0,929	0,049	0,81
High	REFERENCE					
Occupation						
Not working	-0,289	0,012	-0,016	0,941	0,273	0,169
Part-time	-0,301	0,012	-0,078	0,723	0,223	0,279
Full-time	REFERENCE					
Province						
Rest of Netherlands	-0,408	<,001	0,716	<,001	1,124	<,001
South Holland	REFERENCE					
Frequency of Train Use						
At least a few times a month	0,333	0,005	-0,38	0,076	-0,713	<,001
Less often or never	REFERENCE					
Frequency of Oudelijn Use						
At least a few times a month	0,191	0,23	-0,865	0,005	-1,056	<,001
Less often or never	REFERENCE					
Get in/ out at any station*						
In/ out at any station	0,61	<,001	1,208	<,001	0,598	0,017
Never in/out at any station	REFERENCE					
Live/Work near any station*						
Live/work at any station	0,196	0,079	0,018	0,937	-0,178	0,391
Never Live/work at any station	REFERENCE					
Area of Living						
Urban	0,266	0,003	-0,507	0,002	-0,773	<,001
Rural	REFERENCE					
Residency						
Rental property in the private sector	-0,101	0,447	0,542	0,021	0,643	0,003
Rental property in the social rental	-0,234	0,015	0,393	0,03	0,627	<,001
Owner-occupied home	REFERENCE					
Likelihood Ratio Test	<,001					
Pearson	<,001					
Deviance	0,585					
Pseudo R-Square: Cox and Snell	0,133					
Pseudo R-Square: Nagelkerke	0,174					
Pseudo R-Square: McFadden	0,099					
Correct Percentage	72,6%					

*Any station between Leiden Central, The Hague Laan of NOI, Schiedam Centrum, Dordrecht

Within this statement, it can be observed that:

1. **Age:**
   Middle-aged people (35-65) are more likely to (Totally) Disagree than be Neutral (coefficient = 0.789, p = 0.006) or (Totally) Agree (coefficient = 1.048, p < 0.001) with the statement, relative to elderly people (65+). They are also less likely to (Totally) Agree than be neutral with the statement (coefficient = -0.259, p = 0.045). Younger people (under 35) also show the similar pattern, they are more likely to (Totally) Disagree (coefficient = 1.036, p < 0.001) than being neutral or (Totally) Agree (coefficient = 1.437, p < 0.001) with the statement, relative to elderly people (65+). This indicates that younger adults and middle age people are less likely to think this is an important consultation compared to elderly people.
2. **Education Level:**
   People with a low education level are less likely to (Totally) Agree (coefficient = -0.457, p < 0.001) or (Totally) Disagree (coefficient = -0.569, p = 0.013) than be Neutral with the statement compared

to people with a high education level. This result indicates that people from low education level tend to be neutral rather than think that this consultation is important or unimportant for them to give their opinion on.

3. **Occupation:**
   People who are not working are less likely to (Totally) Agree than be Neutral (coefficient = -0.289, p = 0.012) compared to those who work full-time. Similarly, people who work part-time are less likely to (Totally) Agree than be Neutral (coefficient = -0.301, p = 0.012) compared to those who work full-time.
4. **Province:**
   People living in the rest of the Netherlands are less likely to (Totally) Agree (coefficient = -0.408, p < 0.001) and more likely to (Totally) Disagree (coefficient = 0.716, p < 0.001) compared to being Neutral, relative to those living in South Holland. They are also more likely to (Totally) Disagree compared to (Totally) Agree (coefficient = 1.124, p < 0.001). This result indicates that people living far awat from the project tend to think that this consultation is unimportant for them.
5. **Frequency of Train Use:**
   People who travel by train at least a few times a month are more likely to (Totally) Agree than be Neutral (coefficient = 0.333, p = 0.005) and less likely to (Totally) Disagree compared to (Totally) Agree (coefficient = -0.713, p < 0.001) compared to people who travel less frequently by train. This result means that people who travel by train more frequently are more likely to think that this consultation is important for them to give their opinion on.
6. **Frequency of Oude Lijn Use:**
   People who use the Oude Lijn line at least a few times a month are less likely to (Totally) Disagree (coefficient = -0.865, p = 0.005) and less likely to (Totally) Disagree compared to (Totally) Agree (coefficient = -1.056, p < 0.001) compared to those who use it less often or never. This indicates that people who travel through Oude Lijn frequently think this is an important topic for them.
7. **Get in/out at any station:**
   People who ever get in or out at any of the four stations mentioned in the consultation are more likely to (Totally) Agree (coefficient = 0.61, p < 0.001) and more likely to (Totally) Disagree (coefficient = 1.208, p < 0.001) than be Neutral with the statement compared to people who never get in or out at those stations. They are also more likely to (Totally) Disagree compared to (Totally) Agree (coefficient = 0.598, p = 0.017). This tesult indicates that people who ever get in or out at any of the mentioned station are more likely to think that this is unimportant topic for them to give their opinion on.
8. **Area of Living:**
   People living in urban areas are more likely to (Totally) Agree (coefficient = 0.266, p = 0.003) and less likely to (Totally) Disagree (coefficient = -0.507, p = 0.002) compared to being Neutral, relative to those living in rural areas. They are also less likely to (Totally) Disagree compared to (Totally) Agree (coefficient = -0.773, p < 0.001). Therefore, people who live in urban area are more likely to think that this is an important topic for them.
9. **Residency:**
   People who live in rental property in the private sector are more likely to (Totally) Disagree than be Neutral (coefficient = 0.542, p = 0.021) and more likely to (Totally) Disagree compared to (Totally) Agree (coefficient = 0.643, p = 0.003) compared to those who live in an owner-occupied house. People who live in rental property in the social rental sector are more likely to (Totally) Disagree than be Neutral (coefficient = 0.393, p = 0.03) and more likely to (Totally) Disagree compared to (Totally) Agree (coefficient = 0.627, p < 0.001) compared to those who live in an owner-occupied house. This result indicates that people who live in rental houses tend to do not think that this is an important topic for them to give their opinion on, compared to people who live in owner-occupied houses.

The Likelihood Ratio Test was significant (p < 0.001), indicating the model's superiority over the intercept-only model in predicting whether respondents thought it was an important topic to give their opinion on. The Pseudo R-Square values (Cox and Snell, Nagelkerke, McFadden) suggest that the model explains a modest proportion of the variance in the outcome variable. The model correctly classifies 72.6% of

cases.

In conclusion, the analysis of whether respondents thought it was an important topic to give their opinion on reveals several significant insights. Younger and middle-aged people are more likely to (Totally) Disagree than to (Totally) Agree, compared to elderly people. This indicates that younger people tend to think that this topic is not important for them to give their opinion on. Similarly, people who are not working or work part-time are less likely to (Totally) Agree compared to those working full-time, suggesting that they are more likely to think that this topic is unimportant for them.

People with a lower level of education are less likely to have strong opinions (either way) compared to those with higher education, tending to be neutral about whether this topic is important for them. Residents of South Holland, frequent train travelers, and those who use the Oude Lijn tend to think that this is an important topic for them to give their opinion on.

Conversely, people who ever get in or out at certain stations are more likely to think that this is an unimportant topic for them compared to people who do not use these stations. Finally, people living in rental properties, both private and social, are more likely to disagree compared to those in owner-occupied homes, indicating a tendency to think that this is not an important topic to give their opinion on.

Overall, these results suggest that demographic factors such as age, employment status, education level, and living location significantly influence whether respondents think it is important to give their opinion on this topic.

G.2.3 I found the consultation difficult to understand

The third multinomial logistic regression on Oude Lijn dataset was conducted on the statement "*I found the consultation difficult to understand*" and presented presented in Table G.4. The analysis revealed 9 significant attributes when controlled for other independent variables.

Table G.4: Results of multinomial logistic regression for statement: I found the consultation difficult to understand

Attributes	Agree vs Neutral		Disagree vs Neutral		Disagree vs Agree	
	B	Sig.	B	Sig.	B	Sig.
Age						
Younger than 35	0,138	0,529	-0,342	0,013	-0,48	0,015
Between 35 and 65	0,524	0,015	-0,149	0,282	-0,673	<,001
Older than 65	REFERENCE					
Gender						
Man	0,217	0,119	0,004	0,97	-0,213	0,079
Woman	REFERENCE					
Education Level						
Low	0,143	0,456	-0,776	<,001	-0,919	<,001
Medium	-0,136	0,487	-0,396	0,002	-0,259	0,124
High	REFERENCE					
Occupation						
Not working	-0,155	0,377	-0,237	0,05	-0,082	0,592
Part-time	-0,313	0,112	-0,15	0,244	0,164	0,342
Full-time	REFERENCE					
Province						
Rest of Netherlands	0,717	<,001	-0,194	0,051	-0,911	<,001
South Holland	REFERENCE					
Frequency of Train Use						
At least a few times a month	-0,153	0,432	0,455	<,001	0,608	<,001
Less often or never	REFERENCE					
Frequency of Oudelijn Use						
At least a few times a month	0,191	0,403	0,319	0,056	0,128	0,514
Less often or never	REFERENCE					
Get in/ out at any station*						
In/ out at any station	0,916	<,001	-0,448	0,005	-1,365	<,001
Never in/out at any station	REFERENCE					
Live/Work near any station*						
Live/work at any station	0,418	0,013	-0,103	0,371	-0,521	<,001
Never Live/work at any station	REFERENCE					
Area of Living						
Urban	0,23	0,112	0,505	<,001	0,275	0,035
Rural	REFERENCE					
Residency						
Rental property in the private sector	0,344	0,093	0	1	-0,344	0,047
Rental property in the social rental	0,278	0,064	-0,327	0,001	-0,605	<,001
Owner-occupied home	REFERENCE					
Likelihood Ratio Test	<,001					
Pearson	<,001					
Deviance	<,001					
Pseudo R-Square: Cox and Snell	0,107					
Pseudo R-Square: Nagelkerke	0,137					
Pseudo R-Square: McFadden	0,074					
Correct Percentage	72,7%					

**Any station between Leiden Central, The Hague Laan of NOI, Schiedam Centrum, Dordrecht*

Within this statement, it can be observed that:

1. **Age:**
   Younger people (under 35) are less likely to (Totally) Disagree than be Neutral (coefficient = -0.342, p = 0.013) or (Totally) Agree (coefficient = -0.48, p = 0.015) with the statement, relative to elderly people (65+). Middle-aged people (35-65) are more likely to (Totally) Agree than be Neutral (coefficient = 0.524, p = 0.015). They are also less likely to (Totally) Disagree than (Totally) Agree with this statement (coefficient = -0.673, p = <0.001). The results indicate that young adults and middle-aged people tend to perceive this consultation quite difficult to understand, compared to elderly people.
2. **Education Level:**
   People with a low education level are less likely to (Totally) Disagree than be Neutral (coefficient = -0.776, p < 0.001) or (Totally) Agree (coefficient = -0.919, p < 0.001) with the statement compared to people with a high education level. People with a medium education level are less likely to

(Totally) Disagree than be Neutral (coefficient = -0.396, p = 0.002). The results can be interpreted that people with lower and medium education level tend to think that this consultation is difficult to understand.

3. **Occupation:**
   People who are not working are less likely to (Totally) Disagree than be Neutral (coefficient = -0.237, p = 0.050) with the statement compared to those who are working full-time, indicating that they are more likely to be neutral about whether this consultation is difficult to understand.
4. **Province:**
   People living in the rest of the Netherlands are more likely to (Totally) Agree than be Neutral (coefficient = 0.717, p < 0.001) and less likely to (Totally) Disagree than (Totally) Agree (coefficient = -0.911, p < 0.001) with the statement compared to those living in South Holland. This indicates that people who live far from the project tend to think that this is a hard consultation to understand.
5. **Frequency of Train Use:**
   People who travel by train at least a few times a month are more likely to (Totally) Disagree than be Neutral (coefficient = 0.455, p < 0.001) and more likely to (Totally) Disagree than (Totally) Agree (coefficient = 0.608, p < 0.001) with the statement compared to people who travel less frequently by train.
6. **Get in/out at any station:**
   People who get in or out at any station are more likely to (Totally) Agree than be Neutral (coefficient = 0.916, p < 0.001) and less likely to (Totally) Disagree than be Neutral (coefficient = -0.448, p = 0.005) or (Totally) Agree (coefficient = -1.365, p < 0.001) with the statement compared to people who never get in or out at those stations. This result indicates that people who ever get in or out at any of the mentioned stations, tend to think that this consultation is difficutl to understand.
7. **Live/Work near any station:**
   People who live or work near a station are more likely to (Totally) Agree than be Neutral (coefficient = 0.418, p = 0.013) and less likely to (Totally) Disagree than (Totally) Agree (coefficient = -0.521, p < 0.001) with the statement compared to people who do not live or work near any station.
8. **Area of Living:**
   People living in urban areas are more likely to (Totally) Disagree than be Neutral (coefficient = 0.505, p < 0.001) or (Totally) Agree (coefficient = 0.275, p = 0.035) with the statement compared to those living in rural areas, indicating that people who live in an urban area think this consultation is fairly easy to understand.
9. **Residency:**
   People living in rental property in the social rental sector are less likely to (Totally) Disagree than be Neutral (coefficient = -0.327, p = 0.001) or (Totally) Agree (coefficient = -0.605, p < 0.001) with the statement compared to those living in an owner-occupied home.

The Likelihood Ratio Test was significant (p < 0.001), indicating the model's superiority over the intercept-only model in predicting whether respondents found the consultation difficult to understand. The Pseudo R-Square values (Cox and Snell, Nagelkerke, McFadden) suggest that the model explains a modest proportion of the variance in the outcome variable. The model correctly classifies 72.7% of cases.

In conclusion, the analysis of whether respondents found the consultation difficult to understand shows several significant relationships. Younger and middle-aged people are less likely to (Totally) Disagree than be Neutral or (Totally) Agree compared to elderly people, indicating that younger people tend to find the consultation difficult to understand. People with lower education levels are less likely to (Totally) Disagree than be Neutral or (Totally) Agree, suggesting that individuals with lower education levels find the consultation difficult to understand. Additionally, people who are not working are more likely to (Totally) Disagree than be Neutral, indicating that unemployed individuals tend to find the consultation difficult to understand.

People living in the rest of the Netherlands are more likely to (Totally) Agree and less likely to (Totally) Disagree compared to those in South Holland, suggesting that people living closer to the project in South Holland find the consultation easier to understand. Frequent train users are more likely to (Totally) Disagree than be Neutral or (Totally) Agree, indicating that frequent train travelers find the consultation

easier to understand. Conversely, people who get in or out at any station are more likely to (Totally) Agree and less likely to (Totally) Disagree, and those who live or work near a station are more likely to (Totally) Agree and less likely to (Totally) Disagree. This suggests that individuals who use the stations or live/work near them find the consultation difficult to understand.

Urban residents are more likely to (Totally) Disagree than be Neutral or (Totally) Agree, indicating that urban residents find the consultation easier to understand. Lastly, people living in social rental properties are less likely to (Totally) Disagree than be Neutral or (Totally) Agree compared to those in owner-occupied homes, indicating that individuals in owner-occupied homes find the consultation easier to understand compared to those in social rental housing.

Overall, these results suggest that demographic factors such as age, education level, occupation, living location, and frequency of station usage significantly influence whether respondents found the consultation difficult to understand.

G.2.4 By participating in this consultation, I have learned about the choices that the government has to make on this subject

The fourth multinomial logistic regression on Oude Lijn dataset was conducted on the statement "*By participating in this consultation, I have learned about the choices that the government has to make on this subject*" and presented presented in Table G.5. The analysis revealed 8 significant attributes when controlled for other independent variables.

Table G.5: Results of multinomial logistic regression for statement: By participating in this consultation, I have learned about the choices that the government has to make on this subject

Attributes	Agree vs Neutral		Disagree vs Neutral		Disagree vs Agree	
	B	Sig.	B	Sig.	B	Sig.
Age						
Younger than 35	0,221	0,068	0,389	0,145	0,168	0,509
Between 35 and 65	0,128	0,291	0,451	0,092	0,323	0,206
Older than 65	REFERENCE					
Gender						
Man	-0,07	0,387	0,291	0,08	0,362	0,021
Woman	REFERENCE					
Education Level						
Low	-0,061	0,576	-0,594	0,008	-0,533	0,012
Medium	0,028	0,79	-0,661	0,003	-0,689	<,001
High	REFERENCE					
Occupation						
Not working	-0,009	0,93	0,236	0,27	0,245	0,222
Part-time	0,132	0,241	0,159	0,483	0,028	0,896
Full-time	REFERENCE					
Province						
Rest of Netherlands	0,078	0,37	-0,345	0,064	-0,423	0,016
South Holland	REFERENCE					
Frequency of Train Use						
At least a few times a month	0,26	0,024	0,124	0,625	-0,136	0,572
Less often or never	REFERENCE					
Frequency of Oudelijn Use						
At least a few times a month	0,326	0,021	0,952	<,001	0,625	0,019
Less often or never	REFERENCE					
Get in/ out at any station*						
In/ out at any station	-0,317	0,018	-1,198	<,001	-0,88	<,001
Never in/out at any station	REFERENCE					
Live/Work near any station*						
Live/work at any station	-0,041	0,678	0,45	0,019	0,49	0,006
Never Live/work at any station	REFERENCE					
Area of Living						
Urban	0,256	0,003	0,005	0,978	-0,251	0,16
Rural	REFERENCE					
Residency						
Rental property in the private sector	0,171	0,168	0,171	0,48	0	0,998
Rental property in the social rental	-0,113	0,215	0,1	0,606	0,212	0,246
Owner-occupied home	REFERENCE					
Likelihood Ratio Test	<,001					
Pearson	<,001					
Deviance	0,92					
Pseudo R-Square: Cox and Snell	0,031					
Pseudo R-Square: Nagelkerke	0,04					
Pseudo R-Square: McFadden	0,021					
Correct Percentage	70,2%					

*Any station between Leiden Central, The Hague Laan of NOI, Schiedam Centrum, Dordrecht

Within this statement, it can be observed that:

1. **Gender:**
   Men are more likely to (Totally) Disagree than be Neutral (coefficient = 0.291, p = 0.08) or (Totally) Agree (coefficient = 0.362, p = 0.021) compared to women.
2. **Education Level:**
   People with a low education level are less likely to (Totally) Disagree than be Neutral (coefficient = -0.594, p = 0.008) or (Totally) Agree (coefficient = -0.533, p = 0.012) compared to people with a high education level. Similarly, people with a medium education level are less likely to (Totally) Disagree than be Neutral (coefficient = -0.661, p = 0.003) or (Totally) Agree (coefficient = -0.689, p < 0.001) compared to people with a high education level. This result indicates that people with low and medium education level tend to think that they learned something from this consultation.
3. **Province:**

People living in the rest of the Netherlands are less likely to (Totally) Disagree than (Totally) Agree (coefficient = -0.423, p = 0.016) compared to those living in South Holland, suggesting that people who live far from the project learned about the choices the government has to make through this consultation.

4. **Frequency of Train Use:**
   People who travel by train at least a few times a month are more likely to (Totally) Agree than be Neutral (coefficient = 0.26, p = 0.024) compared to people who travel less frequently by train. This indicates that people who travel by train more frequently are more likely to learned the government's choices through this consultation.
5. **Frequency of Oude Lijn Use:**
   People who use the Oude Lijn line at least a few times a month are more likely to (Totally) Agree than be Neutral (coefficient = 0.326, p = 0.021), more likely to (Totally) Disagree than be Neutral (coefficient = 0.952, p < 0.001), and more likely to (Totally) Disagree than (Totally) Agree (coefficient = 0.625, p = 0.019) compared to those who use it less often or never.
6. **Get in/out at any station:**
   People who get in or out at any station are less likely to (Totally) Agree than be Neutral (coefficient = -0.317, p = 0.018), less likely to (Totally) Disagree than be Neutral (coefficient = -1.198, p < 0.001), and less likely to (Totally) Disagree than (Totally) Agree (coefficient = -0.88, p < 0.001) compared to people who never get in or out at those stations.
7. **Live/Work near any station:**
   People who live or work near a station are more likely to (Totally) Disagree than be Neutral (coefficient = 0.45, p = 0.019) and more likely to (Totally) Disagree than (Totally) Agree (coefficient = 0.49, p = 0.006) compared to people who do not live or work near any station.
8. **Area of Living:**
   People living in urban areas are more likely to (Totally) Agree than be Neutral (coefficient = 0.256, p = 0.003) compared to those living in rural areas.

The Likelihood Ratio Test was significant (p < 0.001), indicating the model's superiority over the intercept-only model in predicting whether respondents learned about the choices that the government has to make on this subject. The Pseudo R-Square values (Cox and Snell, Nagelkerke, McFadden) suggest that the model explains a modest proportion of the variance in the outcome variable. The model correctly classifies 70.2% of cases.

In conclusion, the analysis of whether respondents felt that they learned about the choices the government has to make through this consultation reveals several significant insights. Gender appears to influence opinions, with men more likely to (Totally) Disagree than be Neutral or (Totally) Agree compared to women. This suggests that men are less likely to feel that they learned about the government's choices through the consultation.

Education level also plays a crucial role. People with a low education level are less likely to (Totally) Disagree than be Neutral or (Totally) Agree compared to those with a high education level. Similarly, people with a medium education level are less likely to (Totally) Disagree than be Neutral or (Totally) Agree compared to those with a high education level. This indicates that people with low and medium education levels are more likely to feel that they learned something from the consultation.

Geographic location significantly affects opinions as well. People living in the rest of the Netherlands are less likely to (Totally) Disagree than (Totally) Agree compared to those in South Holland, suggesting that individuals living farther from the project feel they learned about the choices the government has to make through the consultation.

Frequency of train use influences perceptions, with people who travel by train at least a few times a month more likely to (Totally) Agree than be Neutral compared to infrequent train travelers. This suggests that frequent train users are more likely to feel they learned about the government's choices through the consultation.

Similarly, frequency of Oude Lijn use also impacts opinions. People who use the Oude Lijn line at least a few times a month are more likely to (Totally) Agree than be Neutral, more likely to (Totally) Disagree

than be Neutral, and more likely to (Totally) Disagree than (Totally) Agree compared to infrequent users. This indicates polarized opinions among frequent Oude Lijn users regarding whether they learned about the government's choices.

Station usage and proximity to stations are also significant factors. People who get in or out at any station are less likely to (Totally) Agree than be Neutral, less likely to (Totally) Disagree than be Neutral, and less likely to (Totally) Disagree than (Totally) Agree compared to those who never use the stations. This suggests a complex view: station users are less inclined to take strong positive or negative stances compared to being Neutral, and among those with strong opinions, they are less likely to hold a negative view compared to a positive one.

People who live or work near a station are more likely to (Totally) Disagree than be Neutral and more likely to (Totally) Disagree than (Totally) Agree compared to those who do not live or work near any station. These findings suggest that station users and residents/workers near stations are less likely to feel that they learned from the consultation.

Lastly, area of living plays a role. People living in urban areas are more likely to (Totally) Agree than be Neutral compared to those in rural areas, indicating that urban residents are more likely to feel that they learned about the choices the government has to make through the consultation.

Overall, these results suggest that demographic factors such as gender, education level, geographic location, and frequency of train and Oude Lijn use significantly influence whether respondents felt they learned about the government's choices through the consultation.

G.2.5 This method should be used more often to involve residents in public policy

The fifth multinomial logistic regression on Oude Lijn dataset was conducted on the statement "*This method should be used more often to involve residents in public policy*" and presented presented in Table G.6. The analysis revealed 10 significant attributes when controlled for other independent variables.

Table G.6: Results of multinomial logistic regression for statement: This method should be used more often to involve residents in public policy

Attributes	Agree vs Neutral		Disagree vs Neutral		Disagree vs Agree	
	B	Sig.	B	Sig.	B	Sig.
Age						
Younger than 35	0,021	0,897	1,581	<,001	1,56	<,001
Between 35 and 65	-0,157	0,316	0,646	0,101	0,802	0,031
Older than 65	REFERENCE					
Gender						
Man	0,086	0,413	0,473	0,03	0,387	0,051
Woman	REFERENCE					
Education Level						
Low	0,001	0,996	-0,014	0,96	-0,015	0,953
Medium	-0,074	0,581	-0,666	0,025	-0,592	0,03
High	REFERENCE					
Occupation						
Not working	-0,158	0,243	0,086	0,75	0,244	0,317
Part-time	-0,159	0,255	0,044	0,879	0,203	0,437
Full-time	REFERENCE					
Province						
Rest of Netherlands	-0,234	0,036	0,339	0,15	0,573	0,008
South Holland	REFERENCE					
Frequency of Train Use						
At least a few times a month	0,184	0,21	-0,342	0,235	-0,526	0,041
Less often or never	REFERENCE					
Frequency of Oudelijn Use						
At least a few times a month	-0,316	0,089	-0,683	0,049	-0,366	0,232
Less often or never	REFERENCE					
Get in/ out at any station*						
In/ out at any station	0,442	0,011	0,609	0,063	0,167	0,566
Never in/out at any station	REFERENCE					
Live/Work near any station*						
Live/work at any station	-0,142	0,261	0,523	0,041	0,665	0,004
Never Live/work at any station	REFERENCE					
Area of Living						
Urban	-0,043	0,707	-1,152	<,001	-1,109	<,001
Rural	REFERENCE					
Residency						
Rental property in the private sector	-0,046	0,771	0,658	0,021	0,703	0,005
Rental property in the social rental	-0,26	0,023	0,204	0,4	0,464	0,037
Owner-occupied home	REFERENCE					
Likelihood Ratio Test	<,001					
Pearson	<,001					
Deviance	1					
Pseudo R-Square: Cox and Snell	0,039					
Pseudo R-Square: Nagelkerke	0,06					
Pseudo R-Square: McFadden	0,038					
Correct Percentage	84,0%					

*Any station between Leiden Central, The Hague Laan of NOI, Schiedam Centrum, Dordrecht

Within this statement, it can be observed that:

1. **Age:**
   Young adults (younger than 35 years old) are more likely to (Totally) Disagree than be Neutral (coefficient = 1.581, $p < 0.001$) or (Totally) Agree (coefficient = 1.56, $p < 0.001$) compared to those 65 and older. People between 35 and 65 are more likely to (Totally) Disagree than (Totally) Agree (coefficient = 0.802, $p = 0.031$) compared to those 65 and older. This suggests that young adults and middle age people tend to think that this method should not be use more often, compared to elderly.
2. **Gender:**
   Men are more likely to (Totally) Disagree than be Neutral (coefficient = 0.473, $p = 0.03$) compared to women, indicating that men are less likely to think this method should be used more often.
3. **Education Level:**

People with a medium education level are less likely to (Totally) Disagree than be Neutral (coefficient = -0.666, p = 0.025) or (Totally) Agree (coefficient = -0.592, p = 0.03) compared to people with a high education level. This result indicates that people from medium education level tend to think this method should be used more often, compared to people from high education background.

4. **Province:**
   People living in the rest of the Netherlands are less likely to (Totally) Agree than be Neutral (coefficient = -0.234, p = 0.036) and more likely to (Totally) Disagree than (Totally) Agree (coefficient = 0.573, p = 0.008) compared to those living in South Holland. It is then indicated that people who live far from the project are less likely to think that this method should be used more often to involve residents.
5. **Frequency of Train Use:**
   People who travel by train at least a few times a month are less likely to (Totally) Disagree than (Totally) Agree (coefficient = -0.526, p = 0.041) compared to people who travel less frequently by train, indicating the tendency for frequent train user to think that this method should be used more often.
6. **Frequency of Oude Lijn Use:**
   People who use the Oude Lijn line at least a few times a month are less likely to (Totally) Disagree than be Neutral (coefficient = -0.683, p = 0.049) compared to those who use it less often or never.
7. **Get in/out at any station:**
   People who get in or out at any station are more likely to (Totally) Agree than be Neutral (coefficient = 0.442, p = 0.011).
8. **Live/Work near any station:**
   People who live or work near a station are more likely to (Totally) Disagree than be Neutral (coefficient = 0.523, p = 0.041) and more likely to (Totally) Disagree than (Totally) Agree (coefficient = 0.665, p = 0.004) compared to people who do not live or work near any station. This results indicate that people who live or work near any of the mentioned station are less likely to think that this method should be used more often.
9. **Area of Living:**
   People living in urban areas are less likely to (Totally) Disagree than be Neutral (coefficient = -1.152, p < 0.001) or (Totally) Agree (coefficient = -1.109, p < 0.001) compared to those living in rural areas, suggesting that urban residents to think that this method should be used more often.
10. **Residency:**
    People living in rental property in the private sector are more likely to (Totally) Disagree than be Neutral (coefficient = 0.658, p = 0.021) or (Totally) Agree (coefficient = 0.703, p = 0.005) compared to those living in owner-occupied homes. People living in rental property in the social rental sector are more likely to (Totally) Disagree than (Totally) Agree (coefficient = 0.464, p = 0.037) compared to those living in owner-occupied homes. This result indicates that people who live in rental properties tend to think that this method should not be used more often, compared to people who live in owner-occupied houses.

The Likelihood Ratio Test was significant (p < 0.001), indicating the model's superiority over the intercept-only model in predicting whether respondents believe this method should be used more often to involve residents in public policy. The Pseudo R-Square values (Cox and Snell, Nagelkerke, McFadden) suggest that the model explains a modest proportion of the variance in the outcome variable. The model correctly classifies 84.0% of cases.

In conclusion, the analysis of whether respondents believe the method should be used more often to involve residents in public policy reveals several significant insights. Age appears to play a crucial role, with young adults (under 35 years old) and middle-aged individuals (35-65 years old) more likely to (Totally) Disagree than be Neutral or (Totally) Agree compared to elderly people. This suggests that younger people tend to think that this method should not be used more often, compared to elderly respondents.

Gender also influences opinions, with men more likely to (Totally) Disagree than be Neutral compared

to women, indicating that men are less likely to support using this method more often. Additionally, people with a medium education level are less likely to (Totally) Disagree than be Neutral or (Totally) Agree compared to those with a high education level, suggesting that those with medium education levels are more supportive of using this method.

Geographic location significantly affects opinions as well. People living in the rest of the Netherlands are less likely to (Totally) Agree and more likely to (Totally) Disagree than those in South Holland, indicating that people living farther from the project are less supportive of using this method more often to involve residents. Furthermore, frequent train users are less likely to (Totally) Disagree than (Totally) Agree compared to infrequent train users, suggesting a tendency for frequent travelers to support using this method more often.

The frequency of Oude Lijn use also plays a role, with frequent users less likely to (Totally) Disagree than be Neutral compared to infrequent users. People who get in or out at any station are more likely to (Totally) Agree than be Neutral, showing support for the method among these respondents. Conversely, people who live or work near a station are more likely to (Totally) Disagree than be Neutral or (Totally) Agree, indicating less support for the method among those living or working near stations.

Urban residents are less likely to (Totally) Disagree than be Neutral or (Totally) Agree compared to rural residents, suggesting that urban dwellers are more supportive of using this method. Lastly, people living in rental properties, both private and social, are more likely to (Totally) Disagree than be Neutral or (Totally) Agree compared to those in owner-occupied homes, indicating that renters are less supportive of using this method more often.

Overall, these results suggest that demographic factors such as age, gender, education level, geographic location, and frequency of train and Oude Lijn use significantly influence whether respondents believe this method should be used more often to involve residents in public policy.

G.2.6 If many people participate in this consultation, then the final decisions on this subject are better accepted for me

The sixth multinomial logistic regression on Oude Lijn dataset was conducted on the statement "*If many people participate in this consultation, then the final decisions on this subject are better accepted for me*" and presented presented in Table G.7. The analysis revealed 5 significant attributes when controlled for other independent variables.

Table G.7: Results of multinomial logistic regression for statement: If many people participate in this consultation, then the final decisions on this subject are better accepted for me

Attributes	Agree vs Neutral		Disagree vs Neutral		Disagree vs Agree	
	B	Sig.	B	Sig.	B	Sig.
Age						
Younger than 35	0,124	0,304	0,732	0,004	0,608	0,012
Between 35 and 65	-0,044	0,715	0,413	0,112	0,457	0,065
Older than 65	REFERENCE					
Gender						
Man	0,054	0,502	0,512	<,001	0,459	0,001
Woman	REFERENCE					
Education Level						
Low	0,195	0,067	-0,216	0,288	-0,411	0,031
Medium	0,31	0,003	-0,16	0,413	-0,47	0,01
High	REFERENCE					
Occupation						
Not working	-0,197	0,058	-0,144	0,46	0,053	0,772
Part-time	-0,204	0,059	-0,116	0,567	0,087	0,646
Full-time	REFERENCE					
Province						
Rest of Netherlands	-0,101	0,241	-0,154	0,358	-0,053	0,735
South Holland	REFERENCE					
Frequency of Train Use						
At least a few times a month	-0,01	0,931	0,057	0,784	0,067	0,731
Less often or never	REFERENCE					
Frequency of Oudelijn Use						
At least a few times a month	0,178	0,196	-0,512	0,036	-0,69	0,002
Less often or never	REFERENCE					
Get in/ out at any station*						
In/ out at any station	-0,046	0,724	0,27	0,259	0,316	0,155
Never in/out at any station	REFERENCE					
Live/Work near any station*						
Live/work at any station	0,027	0,776	0,259	0,154	0,231	0,171
Never Live/work at any station	REFERENCE					
Area of Living						
Urban	-0,022	0,799	-0,622	<,001	-0,600	<,001
Rural	REFERENCE					
Residency						
Rental property in the private sector	-0,005	0,964	0,295	0,148	0,3	0,107
Rental property in the social rental	-0,125	0,165	-0,249	0,18	-0,124	0,48
Owner-occupied home	REFERENCE					
Likelihood Ratio Test	<,001					
Pearson	<,001					
Deviance	0,046					
Pseudo R-Square: Cox and Snell	0,022					
Pseudo R-Square: Nagelkerke	0,028					
Pseudo R-Square: McFadden	0,014					
Correct Percentage	67,9%					

*Any station between Leiden Central, The Hague Laan of NOI, Schiedam Centrum, Dordrecht

Within this statement, it can be observed that:

1. **Age:**
   Younger adults are more likely to (Totally) Disagree than be Neutral (coefficient = 0.732, p = 0.004) and more likely to (Totally) Disagree than (Totally) Agree (coefficient = 0.608, p = 0.012) compared to those 65 and older. This indicates that young adults do not think to better accept the final decisions even if more people participate in this consultation.
2. **Gender:**
   Men are more likely to (Totally) Disagree than be Neutral (coefficient = 0.512, p < 0.001) and more likely to (Totally) Disagree than (Totally) Agree (coefficient = 0.459, p = 0.001) compared to women, indicating men do not better accept the final decisions regarding the Oude Lijn.
3. **Education Level:**
   People with low education level are Less likely to (Totally) Disagree than (Totally) Agree (coeffi-

cient = -0.411, p = 0.031) compared to those with a high education level. Similarly, people with medium education level are also less likely to (Totally) Disagree than (Totally) Agree (coefficient = -0.47, p = 0.01), they are more likely to (Totally) Agree than be Neutral (coefficient = 0.31, p = 0.003) and compared to those with a high education level. This result suggests that people with low and medium education level tend to accept the government's decisions regarding Oude Lijn, if there are many people participated in this consultation.

4. **Frequency of Oude Lijn Use:**
   People who travel through Oude Lijn at least a few times a month are less likely to (Totally) Disagree than be Neutral (coefficient = -0.512, p = 0.036) and less likely to (Totally) Disagree than (Totally) Agree (coefficient = -0.69, p = 0.002) compared to those who use it less often or never. This result indicates that people who frequently travel through Oude Lijn will accept the final decisions of Oude Lijn better.
5. **Area of Living:**
   Urban residents are less likely to (Totally) Disagree than be Neutral (coefficient = -0.622, p < 0.001) and less likely to (Totally) Disagree than (Totally) Agree (coefficient = -0.6, p < 0.001) compared to rural areas. The result suggests that urban dwellers are better at accepting final decisions results, if many people participate in this consultation.

The Likelihood Ratio Test was significant (p < 0.001), indicating the model's superiority over the intercept-only model in predicting whether respondents believe that the final decisions on this subject are better accepted if many people participate in the consultation. The Pseudo R-Square values (Cox and Snell, Nagelkerke, McFadden) suggest that the model explains a modest proportion of the variance in the outcome variable. The model correctly classifies 67.9% of cases.

In conclusion, the analysis of whether respondents believe that if many people participate in this consultation, the final decisions on this subject will be better accepted reveals several significant insights.

Age appears to play a crucial role, with younger adults more likely to (Totally) Disagree than be Neutral or (Totally) Agree compared to those 65 and older. This indicates that younger adults do not think they will better accept the final decisions even if more people participate in this consultation. Similarly, gender influences opinions, with men more likely to (Totally) Disagree than be Neutral or (Totally) Agree compared to women, suggesting that men do not better accept the final decisions regarding the Oude Lijn.

Education level shows significant effects as well. People with a low education level are less likely to (Totally) Disagree than (Totally) Agree compared to those with a high education level. Similarly, people with a medium education level are also less likely to (Totally) Disagree than (Totally) Agree and are more likely to (Totally) Agree than be Neutral compared to those with a high education level. This suggests that people with low and medium education levels tend to accept the government's decisions regarding Oude Lijn if many people participate in this consultation.

Frequency of Oude Lijn use significantly impacts opinions. People who travel through Oude Lijn at least a few times a month are less likely to (Totally) Disagree than be Neutral or (Totally) Agree compared to those who use it less often or never. This indicates that frequent travelers through Oude Lijn are more likely to accept the final decisions.

Area of living also plays a role. Urban residents are less likely to (Totally) Disagree than be Neutral or (Totally) Agree compared to those in rural areas. This suggests that urban dwellers are better at accepting final decisions if many people participate in this consultation.

Overall, these results suggest that demographic factors such as age, gender, education level, frequency of Oude Lijn use, and area of living significantly influence whether respondents believe that the final decisions will be better accepted if many people participate in the consultation. Younger adults and men tend to be less accepting, while individuals with lower and medium education levels, frequent Oude Lijn users, and urban residents are more likely to accept the decisions made following broad participation.

G.2.7 If the government often allows residents to think along about these types of choices more often, then I get more confidence in the decisions of the government

Table G.8: Results of Multinomial logistic regression for statement: If the government often allows residents to think along about these types of choices more often, then I get more confidence in the decisions of the government

Attributes	Agree vs Neutral		Disagree vs Neutral		Disagree vs Agree	
	B	Sig.	B	Sig.	B	Sig.
Age						
Younger than 35	-0,054	0,678	0,610	0,012	0,665	0,003
Between 35 and 65	-0,166	0,203	0,254	0,297	0,420	0,061
Older than 65	REFERENCE					
Gender						
Man	0,026	0,761	0,321	0,035	0,296	0,034
Woman	REFERENCE					
Education Level						
Low	0,126	0,263	0,280	0,154	0,154	0,390
Medium	0,243	0,024	-0,544	0,013	-0,787	<,001
High	REFERENCE					
Occupation						
Not working	0,091	0,404	-0,024	0,903	-0,115	0,522
Part-time	0,029	0,793	-0,003	0,988	-0,032	0,863
Full-time	REFERENCE					
Province						
Rest of Netherlands	0,053	0,563	0,210	0,199	0,157	0,293
South Holland	REFERENCE					
Frequency of Train Use						
At least a few times a month	0,136	0,253	-0,282	0,197	-0,417	0,038
Less often or never	REFERENCE					
Frequency of Oudelijn Use						
At least a few times a month	0,149	0,304	0,498	0,059	0,350	0,147
Less often or never	REFERENCE					
Get in/ out at any station*						
In/ out at any station	0,006	0,965	-0,231	0,345	-0,236	0,288
Never in/out at any station	REFERENCE					
Live/Work near any station*						
Live/work at any station	0,021	0,833	0,107	0,554	0,086	0,604
Never Live/work at any station	REFERENCE					
Area of Living						
Urban	-0,003	0,973	-0,162	0,323	-0,159	0,289
Rural	REFERENCE					
Residency						
Rental property in the private sector	-0,135	0,255	-0,305	0,181	-0,170	0,423
Rental property in the social rental	0,027	0,781	0,224	0,187	0,197	0,201
Owner-occupied home	REFERENCE					
Likelihood Ratio Test	<,001					
Pearson	<,001					
Deviance	0,264					
Pseudo R-Square: Cox and Snell	0,017					
Pseudo R-Square: Nagelkerke	0,022					
Pseudo R-Square: McFadden	0,011					
Correct Percentage	71,3%					

**Any station between Leiden Central, The Hague Laan of NOI, Schiedam Centrum, Dordrecht*

The seventh and last multinomial logistic regression on Oude Lijn dataset was conducted on the statement "*If the government often allows residents to think along about these types of choices more often, then I get more confidence in the decisions of the government*" and presented presented in Table G.8.

The analysis revealed 4 significant attributes when controlled for other independent variables. Within this statement, it can be observed that:

1. **Age:**
   Young adults are more likely to (Totally) Disagree than be Neutral (coefficient = 0.610, p = 0.012) and more likely to (Totally) Disagree than (Totally) Agree (coefficient = 0.665, p = 0.003) compared to those 65 and older. This result suggests that people young adults are less likely to think that they will get more confidence in the governments' decisions.

2. **Gender:**
   Men are more likely to (Totally) Disagree than be Neutral (coefficient = 0.321, p = 0.035) and more likely to (Totally) Disagree than (Totally) Agree (coefficient = 0.296, p = 0.034) compared to women.
3. **Education Level:**
   Medium education level: More likely to (Totally) Agree than be Neutral (coefficient = 0.243, p = 0.024), less likely to (Totally) Disagree than be Neutral (coefficient = -0.544, p = 0.013), and less likely to (Totally) Disagree than (Totally) Agree (coefficient = -0.787, p < 0.001) compared to those with a high education level. This result suggests that people with medium education level tend to think they will get more confidence in the decisions of the governments if residents are allowed to think along about these types of choices more often.
4. **Frequency of Train Use:**
   People who travel by train at least a few times a month are less likely to (Totally) Disagree than (Totally) Agree (coefficient = -0.417, p = 0.038) compared to people who travel less frequently by train.

The Likelihood Ratio Test was significant (p < 0.001), indicating the model's superiority over the intercept-only model in predicting whether respondents gain more confidence in government decisions if residents are often allowed to think along about these types of choices. The Pseudo R-Square values (Cox and Snell, Nagelkerke, McFadden) suggest that the model explains a modest proportion of the variance in the outcome variable. The model correctly classifies 71.3% of cases.

In conclusion, the analysis of whether respondents believe they will gain more confidence in the government's decisions if residents are often allowed to think along about these types of choices reveals several significant insights.

Age appears to play a crucial role, with younger adults more likely to (Totally) Disagree than be Neutral or (Totally) Agree compared to those 65 and older. This result suggests that younger adults are less likely to think that they will gain more confidence in the government's decisions even if residents are more involved in the decision-making process.

Gender also influences opinions, with men more likely to (Totally) Disagree than be Neutral or (Totally) Agree compared to women. This indicates that men are less likely to feel that increased resident involvement will boost their confidence in government decisions.

Education level shows significant effects as well. People with a medium education level are more likely to (Totally) Agree than be Neutral, less likely to (Totally) Disagree than be Neutral, and less likely to (Totally) Disagree than (Totally) Agree compared to those with a high education level. This suggests that individuals with a medium education level tend to believe they will gain more confidence in the government's decisions if residents are allowed to think along about these types of choices more often, compared to people with high education level.

Frequency of train use significantly impacts opinions. People who travel by train at least a few times a month are less likely to (Totally) Disagree than (Totally) Agree compared to those who travel less frequently by train. This indicates that frequent train users are more likely to believe that increased resident involvement will enhance their confidence in government decisions.

Overall, these results suggest that demographic factors such as age, gender, education level, and frequency of train use significantly influence whether respondents believe they will gain more confidence in the government's decisions if residents are often allowed to think along about these types of choices. Younger adults and men tend to be less confident, while individuals with a medium education level and frequent train users are more likely to believe that increased resident involvement will boost their confidence in government decisions.

G.3 Complete Weighted Model Results: Mobility Vision

G.3.1 I trust that this research project is conducted in a fair way

The Multinomial Logistic Regression analysis was conducted for each face validity statement in the Mobility Vision consultation. Similar to previous analysis with Lelylijn and Oude Lijn dataset, the anal-

ysis with Mobility Vision data was conducted with three outcome groups: (Totally) Agree, Neutral, and (Totally) Disagree. The regression was performed twice, using ”Neutral” and ”(Totally) Agree” as the reference outcomes, resulting in three comparisons: (Totally) Agree vs. Neutral, (Totally) Disagree vs. Neutral, and (Totally) Disagree vs. (Totally) Agree.

Table G.9: Results of multinomial logistic regression for statement: I trust that this research project is conducted in a fair way

Attributes	Agree vs Neutral		Disagree vs Neutral		Disagree vs Agree	
	B	Sig.	B	Sig.	B	Sig.
Age						
18-34 years	-0,064	0,689	1,531	<,001	1,595	<,001
35-64 years	-0,283	0,080	0,701	0,113	0,984	0,019
65 years or older	REFERENCE					
Gender						
Man	0,113	0,243	0,912	<,001	0,799	<,001
Woman	REFERENCE					
Education						
Low	-0,414	0,001	0,555	0,040	0,970	<,001
Medium	-0,273	0,016	-0,295	0,279	-0,023	0,929
High	REFERENCE					
Occupation						
Not working	0,008	0,949	-0,064	0,810	-0,072	0,771
Full-time	REFERENCE					
Area of Living						
Urban	-0,058	0,556	-0,008	0,973	0,050	0,811
Rural	REFERENCE					
Frequency of Facing Accessibility Problems						
Low	0,719	<,001	0,315	0,341	-0,404	0,190
Moderate	0,492	0,002	-0,059	0,870	-0,551	0,105
High	REFERENCE					
Likelihood Ratio Test	<,001					
Pearson	<,001					
Deviance	<,001					
Pseudo R-Square: Cox and Snell	0,022					
Pseudo R-Square: Nagelkerke	0,034					
Pseudo R-Square: McFadden	0,022					
Correct Percentage	84,80%					

The ”B” column represents the estimated coefficients for the predictor variables. When the predictors are categorical, the coefficients indicate the change in the log-odds of the outcome categories relative to the reference category of the predictor variable. The ”Sig.” column in the results table indicates the significance level (p-value) for the test of the null hypothesis that the corresponding coefficient is equal to zero. If the p-value is less than 0,05, the null hypothesis is rejected, suggesting that the predictor variable significantly affects the outcome variable. Conversely, if the p-value is greater than or equal to 0,05, there is not enough evidence to suggest a significant effect of the predictor variable on the outcome variable.

The multinomial logistic regression conducted on the statement ”*I trust that this research project is conducted in a fair way*” in the Mobility Vision dataset is presented in Table G.9. The analysis revealed 4 significant attributes when controlled for other independent variables.

Within this statement, it can be observed that:

1. **Age:**
   Individuals aged 18-34 years are more likely to (Totally) Disagree than be Neutral (coefficient = 1,531, $p < 0,001$) and more likely to (Totally) Disagree than (Totally) Agree (coefficient = 1,595, $p < 0,001$) compared to those 65 years or older. This indicates that younger adults are less likely to trust that the research project is conducted fairly compared to seniors. Similarly, individuals aged 35-64 years are more likely to (Totally) Disagree than (Totally) Agree (coefficient = 0,984, $p = 0,019$) compared to those 65 years or older, indicating that middle-aged adults also exhibit less

trust.

2. **Gender:**
   Men are more likely to (Totally) Disagree than be Neutral (coefficient = 0,912, p < 0,001) and more likely to (Totally) Disagree than (Totally) Agree (coefficient = 0,799, p < 0,001) compared to women. This suggests that men are less likely to trust the research project compared to women.
3. **Education Level:**
   People with low education levels are less likely to (Totally) Agree than be Neutral (coefficient = -0,414, p = 0,001) and more likely to (Totally) Disagree than be Neutral (coefficient = 0,555, p = 0,040) and more likely to (Totally) Disagree than (Totally) Agree (coefficient = 0,970, p < 0,001) compared to those with high education levels. This indicates that individuals with low education levels have less trust in the fairness of the research project. Those with medium education levels are less likely to (Totally) Agree than be Neutral (coefficient = -0,273, p = 0,016), suggesting that lower education levels are associated with less trust.
4. **Frequency of Facing Accessibility Problems:**
   People who face low frequency of accessibility problems are more likely to (Totally) Agree than be Neutral (coefficient = 0,719, p < 0,001) and those facing moderate frequency of accessibility problems are also more likely to (Totally) Agree than be Neutral (coefficient = 0,492, p = 0,002) compared to those who face high frequency of accessibility problems, indicating that individuals who face fewer accessibility problems are more likely to trust the research project's fairness.

The Likelihood Ratio Test was significant (p < 0,001), indicating the model's superiority over the intercept-only model in predicting trust in the fairness of the research project. The Pseudo R-Square values (Cox and Snell, Nagelkerke, McFadden) suggest that the model explains a small proportion of the variance in the outcome variable. The model correctly classifies 84,80% of cases.

In conclusion, the analysis of whether respondents trust that the research project is conducted in a fair way reveals several significant insights. Younger individuals and those with lower education levels are less likely to trust the research project's fairness. Men are less likely to trust the project compared to women. Additionally, individuals who face fewer accessibility problems are more likely to trust the project's fairness. These findings suggest that demographic factors such as age, gender, education level, and frequency of facing accessibility problems significantly influence whether respondents trust that the research project is conducted in a fair way.

G.3.2 I thought it was an important topic to give my opinion on

The multinomial logistic regression conducted on the statement "*I thought it was an important topic to give my opinion on*" in the Mobility Vision dataset is presented in Table G.10. The analysis revealed 4 significant attributes when controlled for other independent variables.

Table G.10: Results of multinomial logistic regression for statement: I thought it was an important topic to give my opinion on

Attributes	Agree vs Neutral		Disagree vs Neutral		Disagree vs Agree	
	B	Sig.	B	Sig.	B	Sig.
Age						
18-34 years	-0,300	0,052	0,978	0,023	1,278	0,002
35-64 years	0,071	0,665	0,886	0,048	0,815	0,055
65 years or older	REFERENCE					
Gender						
Man	0,470	<,001	0,834	<,001	0,364	0,077
Woman	REFERENCE					
Education						
Low	-0,730	<,001	0,880	0,003	1,610	<,001
Medium	-0,433	<,001	0,323	0,269	0,757	0,006
High	REFERENCE					
Occupation						
Not working	0,038	0,747	-0,279	0,304	-0,317	0,210
Full-time	REFERENCE					
Area of Living						
Urban	-0,057	0,563	0,263	0,261	0,321	0,143
Rural	REFERENCE					
Frequency of Facing Accessibility Problems						
Low	-0,107	0,545	-0,869	0,006	-0,762	0,007
Moderate	0,011	0,952	-0,693	0,047	-0,704	0,023
High	REFERENCE					
Likelihood Ratio Test	<,001					
Pearson	<,001					
Deviance	<,001					
Pseudo R-Square: Cox and Snell	0,033					
Pseudo R-Square: Nagelkerke	0,052					
Pseudo R-Square: McFadden	0,034					
Correct Percentage	85,00%					

The Likelihood Ratio Test was significant ($p < 0,001$), indicating the model's superiority over the intercept-only model in predicting whether respondents thought the topic was important. The Pseudo R-Square values (Cox and Snell, Nagelkerke, McFadden) suggest that the model explains a small proportion of the variance in the outcome variable. The model correctly classifies 85,00% of cases. However, the Pearson and Deviance statistics both have p-values less than 0,001, indicating that the model does not fit the data perfectly, and there may be some discrepancies between the observed and expected frequencies.

Despite the significant Pearson and Deviance statistics suggesting a less than perfect fit, the model's ability to correctly classify 85,00% of cases and the significant Likelihood Ratio Test indicate that the model still has substantial predictive power.

Within this statement, it can be observed that:

1. **Age:**
   Individuals aged 18-34 years are more likely to (Totally) Disagree than be Neutral ($B = 0,978$, $p = 0,023$) and more likely to (Totally) Disagree than (Totally) Agree ($B = 1,278$, $p = 0,002$) compared to those 65 years or older. This suggests that younger adults are less likely to consider the topic important compared to seniors. Similarly, individuals aged 35-64 years are more likely to (Totally) Disagree than be Neutral ($B = 0,886$, $p = 0,048$) compared to those 65 years or older, indicating that middle-aged adults also find the topic less important.
2. **Gender:**
   Men are more likely to (Totally) Agree than be Neutral ($B = 0,470$, $p < 0,001$) and more likely to (Totally) Disagree than be Neutral ($B = 0,834$, $p < 0,001$) compared to women. This suggests that men are more polarized in their opinions, either agreeing or disagreeing strongly compared to women.

3. **Education Level:**
   Individuals with low education levels are less likely to (Totally) Agree than be Neutral (B = -0,730, p < 0,001) and more likely to (Totally) Disagree than be Neutral (B = 0,880, p = 0,003) and more likely to (Totally) Disagree than (Totally) Agree (B = 1,610, p < 0,001) compared to those with high education levels. This indicates that individuals with low education levels do not consider the topic as important. Those with medium education levels are less likely to (Totally) Agree than be Neutral (B = -0,433, p < 0,001) and more likely to (Totally) Disagree than (Totally) Agree (B = 0,757, p = 0,006), suggesting a similar trend with low education level individuals.
4. **Frequency of Facing Accessibility Problems:**
   Individuals who face low frequency of accessibility problems are less likely to (Totally) Disagree than be Neutral (B = -0,869, p = 0,006) and less likely to (Totally) Disagree than (Totally) Agree (B = -0,762, p = 0,007) compared to those who face high frequency of accessibility problems. This indicates that individuals who face high frequency of accessibility problems consider the topic more important. Similarly, those facing moderate frequency of accessibility problems are less likely to (Totally) Disagree than be Neutral (B = -0,693, p = 0,047) and less likely to (Totally) Disagree than (Totally) Agree (B = -0,704, p = 0,023), suggesting the same trend.

In conclusion, the analysis of whether respondents thought it was an important topic to give their opinion on reveals several significant insights. Younger individuals and those with lower education levels are less likely to consider the topic important. Men show more polarized opinions compared to women. Additionally, individuals who face fewer accessibility problems are more likely to consider the topic important. These findings suggest that demographic factors such as age, gender, education level, and frequency of facing accessibility problems significantly influence whether respondents think the topic is important.

Based on the provided Multinomial Logistic Regression (MLR) results for the statement ”I found the consultation difficult to understand” in the Mobility Vision consultation, here is the interpretation in LaTeX format:

G.3.3 I found the consultation difficult to understand

The multinomial logistic regression conducted on the statement ”*I found the consultation difficult to understand*” in the Mobility Vision dataset is presented in Table G.11. The analysis revealed several significant attributes when controlled for other independent variables.

The Likelihood Ratio Test was significant (p < 0,001), indicating the model’s superiority over the intercept-only model in predicting whether respondents found the consultation difficult to understand. The Pseudo R-Square values (Cox and Snell, Nagelkerke, McFadden) suggest that the model explains a small proportion of the variance in the outcome variable. The model correctly classifies 72,70% of cases. However, the Pearson and Deviance statistics both have p-values less than 0,001, indicating that the model does not fit the data perfectly, and there may be some discrepancies between the observed and expected frequencies.

Despite the significant Pearson and Deviance statistics suggesting a less than perfect fit, the model’s ability to correctly classify 72,70% of cases and the significant Likelihood Ratio Test indicate that the model still has substantial predictive power.

Table G.11: Results of multinomial logistic regression for statement: I found the consultation difficult to understand

Attributes	Agree vs Neutral		Disagree vs Neutral		Disagree vs Agree	
	B	Sig.	B	Sig.	B	Sig.
Age						
18-34 years	1,157	<,001	-0,194	0,155	-1,351	<,001
35-64 years	0,670	0,007	0,260	0,070	-0,409	0,065
65 years or older	REFERENCE					
Gender						
Man	0,865	<,001	0,333	<,001	-0,532	<,001
Woman	REFERENCE					
Education						
Low	0,037	0,828	-1,564	<,001	-1,601	<,001
Medium	-0,183	0,266	-0,761	<,001	-0,579	<,001
High	REFERENCE					
Occupation						
Not working	-0,235	0,143	0,272	0,012	0,506	<,001
Full-time	REFERENCE					
Area of Living						
Urban	0,375	0,006	-0,045	0,611	-0,420	<,001
Rural	REFERENCE					
Frequency of Facing Accessibility Problems						
Low	0,143	0,525	0,075	0,629	-0,068	0,728
Moderate	0,004	0,987	-0,152	0,356	-0,156	0,449
High	REFERENCE					
Likelihood Ratio Test	<,001					
Pearson	<,001					
Deviance	<,001					
Pseudo R-Square: Cox and Snell	0,098					
Pseudo R-Square: Nagelkerke	0,125					
Pseudo R-Square: McFadden	0,067					
Correct Percentage	72,70%					

Within this statement, it can be observed that:

1. **Age:**
   Individuals aged 18-34 years are more likely to (Totally) Agree than be Neutral (B = 1,157, p < 0,001) and less likely to (Totally) Disagree than (Totally) Agree (B = -1,351, p < 0,001) compared to those 65 years or older. This suggests that younger adults find the consultation more difficult to understand. Similarly, individuals aged 35-64 years are more likely to (Totally) Agree than be Neutral (B = 0,670, p = 0,007) compared to those 65 years or older, indicating that middle-aged adults also find the consultation more difficult to understand.
2. **Gender:**
   Men are more likely to (Totally) Agree than be Neutral (B = 0,865, p < 0,001) and more likely to (Totally) Disagree than be Neutral (B = 0,333, p < 0,001) and less likely to (Totally) Disagree than (Totally) Agree (B = -0,532, p < 0,001) compared to women. This suggests that men find the consultation more difficult to understand compared to women.
3. **Education Level:**
   Individuals with low education levels are less likely to (Totally) Disagree than be Neutral (B = -1,564, p < 0,001) and less likely to (Totally) Disagree than (Totally) Agree (B = -1,601, p < 0,001) compared to those with high education levels. Similarly, those with medium education levels are also less likely to (Totally) Disagree than be Neutral (B = -0,761, p < 0,001) and less likely to (Totally) Disagree than (Totally) Agree (B = -0,579, p < 0,001). This indicates that individuals with low and medium education levels find the consultation more difficult to understand compared to highly educated individuals.
4. **Occupation:**
   Individuals who are not working are more likely to (Totally) Disagree than be Neutral (B = 0,272, p = 0,012) and more likely to (Totally) Disagree than (Totally) Agree (B = 0,506, p < 0,001) compared

to those working full-time, indicating that unemployed individuals find the consultation less difficult to understand.

5. **Area of Living:**
   Individuals living in urban areas are more likely to (Totally) Agree than be Neutral (B = 0,375, p = 0,006) and less likely to (Totally) Disagree than (Totally) Agree (B = -0,420, p < 0,001) compared to those living in rural areas, indicating that urban residents find the consultation more difficult to understand.

The frequency of facing accessibility problems did not show significant results in this analysis.

In conclusion, the analysis of whether respondents found the consultation difficult to understand reveals several significant insights. Younger individuals and men are more likely to find the consultation difficult to understand. Additionally, individuals with lower education levels and those not working find the consultation more difficult to understand. Urban residents are more likely to find the consultation difficult to understand compared to rural residents. These findings suggest that demographic factors such as age, gender, education level, occupation, and area of living significantly influence whether respondents find the consultation difficult to understand.

G.3.4 This method should be used more often to involve residents in government policies

The multinomial logistic regression conducted on the statement "*This method should be used more often to involve residents in government policies*" in the Mobility Vision dataset is presented in Table G.12. The analysis revealed 5 significant attributes when controlled for other independent variables.

The Likelihood Ratio Test was significant (p < 0,001), indicating the model's superiority over the intercept-only model in predicting whether respondents agree that this method should be used more often. The Pseudo R-Square values (Cox and Snell, Nagelkerke, McFadden) suggest that the model explains a small proportion of the variance in the outcome variable. The model correctly classifies 79,30% of cases. However, the Pearson and Deviance statistics both have p-values less than 0,001, indicating that the model does not fit the data perfectly, and there may be some discrepancies between the observed and expected frequencies.

Table G.12: Results of multinomial logistic regression for statement: This method should be used more often to involve residents in government policies

Attributes	Agree vs Neutral		Disagree vs Neutral		Disagree vs Agree	
	B	Sig.	B	Sig.	B	Sig.
Age						
18-34 years	0,031	0,815	0,076	0,828	0,045	0,893
35-64 years	0,373	0,008	0,289	0,422	-0,084	0,806
65 years or older	REFERENCE					
Gender						
Man	0,283	<,001	1,080	<,001	0,797	<,001
Woman	REFERENCE					
Education						
Low	-0,368	<,001	0,098	0,690	0,466	0,045
Medium	0,027	0,785	-0,203	0,395	-0,231	0,304
High	REFERENCE					
Occupation						
Not working	0,159	0,134	-0,507	0,076	-0,666	0,015
Full-time	REFERENCE					
Area of Living						
Urban	-0,085	0,324	0,237	0,272	0,322	0,114
Rural	REFERENCE					
Frequency of Facing Accessibility Problems						
Low	-0,182	0,254	-0,747	0,014	-0,565	0,039
Moderate	-0,303	0,069	-0,866	0,009	-0,563	0,062
High	REFERENCE					
Likelihood Ratio Test	<,001					
Pearson	<,001					
Deviance	<,001					
Pseudo R-Square: Cox and Snell	0,021					
Pseudo R-Square: Nagelkerke	0,031					
Pseudo R-Square: McFadden	0,018					
Correct Percentage	79,30%					

Despite the significant Pearson and Deviance statistics suggesting a less than perfect fit, the model's ability to correctly classify 79,30% of cases and the significant Likelihood Ratio Test indicate that the model still has substantial predictive power.

Within this statement, it can be observed that:

1. **Age:**
   Individuals aged 35-64 years are more likely to (Totally) Agree than be Neutral (B = 0,373, p = 0,008) compared to those 65 years or older. This suggests that middle-aged adults are more likely to think that this method should be used more often to involve residents in government policies. The differences between other age groups were not significant.
2. **Gender:**
   Men are more likely to (Totally) Agree than be Neutral (B = 0,283, p < 0,001), more likely to (Totally) Disagree than be Neutral (B = 1,080, p < 0,001), and more likely to (Totally) Disagree than (Totally) Agree (B = 0,797, p < 0,001) compared to women. This suggests that men are more likely to disagree with the statement.
3. **Education Level:**
   Individuals with low education levels are less likely to (Totally) Agree than be Neutral (B = -0,368, p < 0,001) and more likely to (Totally) Disagree than (Totally) Agree (B = 0,466, p = 0,045) compared to those with high education levels. This indicates that individuals with lower education levels are less supportive of using this method more often to involve residents in government policies.
4. **Occupation:**
   Individuals who are not working are less likely to (Totally) Disagree than (Totally) Agree (B = -0,666, p = 0,015) compared to those working full-time, indicating that unemployed individuals are more supportive of using this method more often.

5. **Frequency of Facing Accessibility Problems:**
   Individuals who face low frequency of accessibility problems are less likely to (Totally) Disagree than be Neutral (B = -0,747, p = 0,014) and less likely to (Totally) Disagree than (Totally) Agree (B = -0,565, p = 0,039) compared to those who face high frequency of accessibility problems. Similarly, those facing moderate frequency of accessibility problems are less likely to (Totally) Disagree than be Neutral (B = -0,866, p = 0,009) compared to those facing high frequency of accessibility problems. This suggests that individuals who face fewer accessibility problems are more supportive of using this method more often.

In conclusion, the analysis of whether respondents believe that this method should be used more often to involve residents in government policies reveals several significant insights. Middle-aged individuals are more supportive of this method. Men are more likely to disagree with the statement. Individuals with lower education levels are less supportive, while unemployed individuals are more supportive. Additionally, individuals who face fewer accessibility problems are more likely to support using this method more often. These findings suggest that demographic factors such as age, gender, education level, occupation, and frequency of facing accessibility problems significantly influence whether respondents think this method should be used more often.

G.3.5 By participating in this consultation, I learned about the choices the Province must make on this topic

The multinomial logistic regression conducted on the statement "*By participating in this consultation, I learned about the choices the Province must make on this topic*" in the Mobility Vision dataset is presented in Table G.13. The analysis revealed 4 significant attributes when controlled for other independent variables.

The Likelihood Ratio Test was significant ($p < 0,001$), indicating the model's superiority over the intercept-only model in predicting whether respondents agreed that they learned about the choices the Province must make. The Pseudo R-Square values (Cox and Snell, Nagelkerke, McFadden) suggest that the model explains a small proportion of the variance in the outcome variable. The model correctly classifies 61,40% of cases. However, the Pearson and Deviance statistics both have p-values less than 0,001, indicating that the model does not fit the data perfectly, and there may be some discrepancies between the observed and expected frequencies.

Despite the significant Pearson and Deviance statistics suggesting a less than perfect fit, the model's ability to correctly classify 61,40% of cases and the significant Likelihood Ratio Test indicate that the model still has substantial predictive power.

Table G.13: Results of multinomial logistic regression for statement: By participating in this consultation, I learned about the choices the Province must make on this topic

Attributes	Agree vs Neutral		Disagree vs Neutral		Disagree vs Agree	
	B	Sig.	B	Sig.	B	Sig.
Age						
18-34 years	0,317	0,005	0,708	<,001	0,391	0,046
35-64 years	0,085	0,465	0,571	0,008	0,486	0,017
65 years or older	REFERENCE					
Gender						
Man	0,086	0,226	0,477	<,001	0,391	<,001
Woman	REFERENCE					
Education						
Low	0,345	<,001	-0,631	<,001	-0,976	<,001
Medium	0,214	0,009	-0,731	<,001	-0,945	<,001
High	REFERENCE					
Occupation						
Not working	-0,094	0,301	0,134	0,390	0,228	0,121
Full-time	REFERENCE					
Area of Living						
Urban	-0,015	0,834	0,123	0,340	0,138	0,256
Rural	REFERENCE					
Frequency of Facing Accessibility Problems						
Low	0,277	0,027	0,069	0,741	-0,209	0,290
Moderate	0,172	0,196	0,076	0,727	-0,096	0,641
High	REFERENCE					
Likelihood Ratio Test	<,001					
Pearson	<,001					
Deviance	<,001					
Pseudo R-Square: Cox and Snell	0,03					
Pseudo R-Square: Nagelkerke	0,036					
Pseudo R-Square: McFadden	0,017					
Correct Percentage	61,40%					

Within this statement, it can be observed that:

1. **Age:**
   Individuals aged 18-34 years are more likely to (Totally) Agree than be Neutral (B = 0,317, p = 0,005), more likely to (Totally) Disagree than be Neutral (B = 0,708, p < 0,001), and more likely to (Totally) Disagree than (Totally) Agree (B = 0,391, p = 0,046) compared to those 65 years or older. This suggests that younger adults are more likely to feel that they did not learn about the choices the Province must make. Similarly, individuals aged 35-64 years are more likely to (Totally) Disagree than be Neutral (B = 0,571, p = 0,008) and more likely to (Totally) Disagree than (Totally) Agree (B = 0,486, p = 0,017) compared to those 65 years or older, indicating a similar trend.
2. **Gender:**
   Men are more likely to (Totally) Disagree than be Neutral (B = 0,477, p < 0,001) and more likely to (Totally) Disagree than (Totally) Agree (B = 0,391, p < 0,001) compared to women. This suggests that men are more likely to feel that they did not learn about the choices the Province must make through the consultation.
3. **Education Level:**
   Individuals with low education levels are more likely to (Totally) Agree than be Neutral (B = 0,345, p < 0,001), less likely to (Totally) Disagree than be Neutral (B = -0,631, p < 0,001), and less likely to (Totally) Disagree than (Totally) Agree (B = -0,976, p < 0,001) compared to those with high education levels. This indicates that individuals with lower education levels are more likely to feel that they learned about the choices the Province must make. Those with medium education levels are also more likely to (Totally) Agree than be Neutral (B = 0,214, p = 0,009), less likely to (Totally) Disagree than be Neutral (B = -0,731, p < 0,001), and less likely to (Totally) Disagree than (Totally) Agree (B = -0,945, p < 0,001), suggesting a similar trend.

4. **Frequency of Facing Accessibility Problems:**
   Individuals who face low frequency of accessibility problems are more likely to (Totally) Agree than be Neutral (B = 0,277, p = 0,027) compared to those who face high frequency of accessibility problems. This suggests that individuals with fewer accessibility problems are more likely to feel that they learned about the choices the Province must make through the consultation.

In conclusion, the analysis of whether respondents believe that they learned about the choices the Province must make by participating in the consultation reveals several significant insights. Younger individuals and men are more likely to feel that they did not learn about the choices the Province must make. Individuals with lower education levels are more likely to feel that they learned about the choices, as are those facing fewer accessibility problems. These findings suggest that demographic factors such as age, gender, education level, and frequency of facing accessibility problems significantly influence whether respondents think they learned about the choices the Province must make.

G.3.6 If many people participate in this consultation, the final decisions on this topic are more acceptable to me

The multinomial logistic regression conducted on the statement "*If many people participate in this consultation, the final decisions on this topic are more acceptable to me*" in the Mobility Vision dataset is presented in Table G.14. The analysis revealed 5 significant attributes when controlled for other independent variables.

Table G.14: Results of multinomial logistic regression for statement: If many people participate in this consultation, the final decisions on this topic are more acceptable to me

Attributes	Agree vs Neutral		Disagree vs Neutral		Disagree vs Agree	
	B	Sig.	B	Sig.	B	Sig.
Age						
18-34 years	0,206	0,066	0,739	<,001	0,533	0,013
35-64 years	-0,027	0,816	0,444	0,056	0,471	0,034
65 years or older	REFERENCE					
Gender						
Man	0,423	<,001	0,966	<,001	0,543	<,001
Woman	REFERENCE					
Education						
Low	0,240	0,010	-0,417	0,017	-0,657	<,001
Medium	0,074	0,360	-0,466	0,001	-0,540	<,001
High	REFERENCE					
Occupation						
Not working	-0,148	0,095	-0,249	0,133	-0,102	0,520
Full-time	REFERENCE					
Area of Living						
Urban	-0,036	0,615	0,263	0,053	0,299	0,020
Rural	REFERENCE					
Frequency of Facing Accessibility Problems						
Low	-0,014	0,912	-0,443	0,026	-0,429	0,018
Moderate	-0,103	0,454	-0,590	0,006	-0,487	0,013
High	REFERENCE					
Likelihood Ratio Test	<,001					
Pearson	<,001					
Deviance	<,001					
Pseudo R-Square: Cox and Snell	0,035					
Pseudo R-Square: Nagelkerke	0,043					
Pseudo R-Square: McFadden	0,021					
Correct Percentage	61,00%					

The Likelihood Ratio Test was significant ($p < 0,001$), indicating the model's superiority over the intercept-only model in predicting whether respondents agreed that the final decisions would be more acceptable with greater participation. The Pseudo R-Square values (Cox and Snell, Nagelkerke, McFadden) suggest that the model explains a small proportion of the variance in the outcome variable. The model correctly classifies 61,00% of cases. However, the Pearson and Deviance statistics both have p-values

less than 0,001, indicating that the model does not fit the data perfectly, and there may be some discrepancies between the observed and expected frequencies.

Despite the significant Pearson and Deviance statistics suggesting a less than perfect fit, the model's ability to correctly classify 61,00% of cases and the significant Likelihood Ratio Test indicate that the model still has substantial predictive power.

Within this statement, it can be observed that:

1. **Age:**
   Individuals aged 18-34 years are more likely to (Totally) Disagree than be Neutral (B = 0,739, p < 0,001), and more likely to (Totally) Disagree than (Totally) Agree (B = 0,533, p = 0,013) compared to those 65 years or older. Similarly, individuals aged 35-64 years are more likely to (Totally) Disagree than be Neutral (B = 0,444, p = 0,056) and more likely to (Totally) Disagree than (Totally) Agree (B = 0,471, p = 0,034) compared to those 65 years or older. This suggests that younger adults and middle aged individuals are more likely to feel that the final decisions would not be more acceptable with greater participation.
2. **Gender:**
   Men are more likely to (Totally) Agree than be Neutral (B = 0,423, p < 0,001), more likely to (Totally) Disagree than be Neutral (B = 0,966, p < 0,001), and more likely to (Totally) Disagree than (Totally) Agree (B = 0,543, p < 0,001) compared to women. This suggests that men are more likely to feel that the final decisions would not be more acceptable with greater participation.
3. **Education Level:**
   Individuals with low education levels are more likely to (Totally) Agree than be Neutral (B = 0,240, p = 0,010), less likely to (Totally) Disagree than be Neutral (B = -0,417, p = 0,017), and less likely to (Totally) Disagree than (Totally) Agree (B = -0,657, p < 0,001) compared to those with high education levels. Those with medium education levels are also less likely to (Totally) Disagree than be Neutral (B = -0,466, p = 0,001) and less likely to (Totally) Disagree than (Totally) Agree (B = -0,540, p < 0,001). This indicates that individuals with lower and medium education levels are more likely to feel that the final decisions would be more acceptable with greater participation, compared to individuals with higher education levels.
4. **Area of Living:**
   Individuals living in urban areas are more likely to (Totally) Disagree than (Totally) Agree (B = 0,299, p = 0,020) compared to those living in rural areas. This suggests that urban residents are more likely to feel that the final decisions would not be more acceptable with greater participation.
5. **Frequency of Facing Accessibility Problems:**
   Individuals who face low frequency of accessibility problems are less likely to (Totally) Disagree than be Neutral (B = -0,443, p = 0,026) and less likely to (Totally) Disagree than (Totally) Agree (B = -0,429, p = 0,018) compared to those who face high frequency of accessibility problems. Similarly, those facing moderate frequency of accessibility problems are less likely to (Totally) Disagree than be Neutral (B = -0,590, p = 0,006) and less likely to (Totally) Disagree than (Totally) Agree (B = -0,487, p = 0,013) compared to those facing high frequency of accessibility problems. This suggests that individuals who face fewer accessibility problems are more likely to feel that the final decisions would be more acceptable with greater participation, compared to people who face problems in accessibility frequently.

In conclusion, the analysis of whether respondents believe that the final decisions on this topic would be more acceptable with greater participation reveals several significant insights. Younger individuals and men are more likely to feel that the final decisions would not be more acceptable with greater participation. Individuals with lower education levels and those facing fewer accessibility problems are more likely to feel that the final decisions would be more acceptable with greater participation. These findings suggest that demographic factors such as age, gender, education level, area of living, and frequency of facing accessibility problems significantly influence whether respondents think the final decisions would be more acceptable with greater participation.

G.3.7 If the government involves residents more often in thinking about these kinds of choices, I will have more confidence in the government's decisions

The multinomial logistic regression conducted on the statement "*If the government involves residents more often in thinking about these kinds of choices, I will have more confidence in the government's decisions*" in the Mobility Vision dataset is presented in Table G.15. The analysis revealed 3 significant attributes when controlled for other independent variables.

Table G.15: Results of multinomial logistic regression for statement: If the government involves residents more often in thinking about these kinds of choices, I will have more confidence in the government's decisions

Attributes	Agree vs Neutral		Disagree vs Neutral		Disagree vs Agree	
	B	Sig.	B	Sig.	B	Sig.
Age						
18-34 years	0,160	0,167	0,112	0,617	-0,049	0,817
35-64 years	0,118	0,327	0,250	0,277	0,132	0,544
65 years or older	REFERENCE					
Gender						
Man	0,352	<,001	0,445	<,001	0,093	0,452
Woman	REFERENCE					
Education						
Low	0,210	0,028	-0,611	0,001	-0,821	<,001
Medium	0,227	0,007	-0,228	0,121	-0,455	<,001
High	REFERENCE					
Occupation						
Not working	0,014	0,881	-0,157	0,369	-0,171	0,298
Full-time	REFERENCE					
Area of Living						
Urban	-0,040	0,594	0,195	0,166	0,234	0,075
Rural	REFERENCE					
Frequency of Facing Accessibility Problems						
Low	-0,133	0,315	0,302	0,253	0,434	0,080
Moderate	-0,113	0,418	0,416	0,129	0,529	0,040
High	REFERENCE					
Likelihood Ratio Test	<,001					
Pearson	<,001					
Deviance	<,001					
Pseudo R-Square: Cox and Snell	0,018					
Pseudo R-Square: Nagelkerke	0,023					
Pseudo R-Square: McFadden	0,011					
Correct Percentage	65,70%					

The Likelihood Ratio Test was significant ($p < 0,001$), indicating the model's superiority over the intercept-only model in predicting whether respondents agreed that involving residents more often would increase their confidence in the government's decisions. The Pseudo R-Square values (Cox and Snell, Nagelkerke, McFadden) suggest that the model explains a small proportion of the variance in the outcome variable. The model correctly classifies 65,70% of cases. However, the Pearson and Deviance statistics both have p-values less than 0,001, indicating that the model does not fit the data perfectly, and there may be some discrepancies between the observed and expected frequencies.

Despite the significant Pearson and Deviance statistics suggesting a less than perfect fit, the model's ability to correctly classify 65,70% of cases and the significant Likelihood Ratio Test indicate that the model still has substantial predictive power.

Within this statement, it can be observed that:

1. **Gender:**
   Men are more likely to (Totally) Agree than be Neutral ($B = 0,352$, $p < 0,001$) and more likely to (Totally) Disagree than be Neutral ($B = 0,445$, $p < 0,001$) compared to women. This suggests that men have more polarized opinions regarding whether involving residents more often would increase their confidence in the government's decisions, with some strongly agreeing and others strongly disagreeing.

2. **Education Level:**
   Individuals with low education levels are more likely to (Totally) Agree than be Neutral (B = 0,210, p = 0,028), less likely to (Totally) Disagree than be Neutral (B = -0,611, p = 0,001), and less likely to (Totally) Disagree than (Totally) Agree (B = -0,821, p < 0,001) compared to those with high education levels. Similarly, those with medium education levels are also more likely to (Totally) Agree than be Neutral (B = 0,227, p = 0,007) and less likely to (Totally) Disagree than (Totally) Agree (B = -0,455, p < 0,001). This indicates that individuals with lower and medium education levels are more likely to feel that involving residents more often would increase their confidence in the government's decisions.
3. **Frequency of Facing Accessibility Problems:**
   Individuals who face moderate frequency of accessibility problems are less likely to (Totally) Disagree than (Totally) Agree (B = 0,529, p = 0,040) compared to those who face high frequency of accessibility problems. This suggests that individuals who face moderate accessibility problems are more likely to feel that involving residents more often would increase their confidence in the government's decisions.

In conclusion, the analysis of whether respondents believe that involving residents more often in thinking about these kinds of choices would increase their confidence in the government's decisions reveals several significant insights. Men have more polarized opinions, with some strongly agreeing and others strongly disagreeing. Individuals with lower education levels and those facing moderate accessibility problems are more likely to feel that involving residents more often would increase their confidence in the government's decisions. These findings suggest that demographic factors such as gender, education level, and frequency of facing accessibility problems significantly influence whether respondents think involving residents more often would increase their confidence in the government's decisions.

H Complete Results of Content Analysis

This chapter provides the detailed percentage per respondents level of each face validity dimensions comments about the dislikes of Lelylijn consultation.

Table H.1: Results of Content Analysis: Why Respondents Dislike the Consultation

Face Validity Dimension	Category	Statement Example
Acceptance	Online form of consultation	Because it is online, you do not reach large parts of the residents who could benefit from the Lely Line (the elderly, people who are less online, who do not understand these types of questionnaires). So you will get a distorted picture anyway.
	Questions lack of depth	Few SMART questions, so the open answers can go in any direction.
	Limited chance of giving opinion	That some questions do not leave room to answer according to your own interpretation. There is room for this occasionally, but it should be possible more often.
	Government/ Experts must take action	The government should not be guided too much by ignorant residents, but by scientific substantiation from experts.
	Focused on Randstad / Northern area only	Opinions don't matter. Government does what they want anyway.. No more grieving in politics
	Preference to invest in other projects	Focusing on the Lely line alone, no alternatives.
Clarity	Unclear in general	Didn't describe everything clearly
	Unclear questions/ choices	Too complicated and too long. Unclear questions. I've done better consultations.
Completeness	Minimum information on the follow up of this consultation	I remain curious and also a bit suspicious about what will be done with this consultation. It would be nice if you provided feedback about this research by e-mail.
	Minimum information about the project	No concrete information about how and what exactly the Lely line will entail in terms of benefits
	Minimum information about the project's impact	Very little prior information. Can 'experts' say anything about environmental effects, economic effects, demographic effects? If so, that influences choices. Does such a consultation help decision-making? I hope that arguments will soon count.
	Minimum information about what experts' think	Get more information from independent experts or from experts who are for and against to form a better opinion.
Feasibility	Confusing choice experiments	I found the part with choices between 2 proposals difficult to understand. Something with sliders regarding ideal characteristics would be easier. Now actually all the proposals were not really what I
	Consultation is too long	A bit too extensive. I think a lot of people will drop out when they realize how much time it takes.
	Points allocation is confusing	Some choices are clustered too much, making the distribution of points very difficult. A quick page with an overview of what you have completed as most important would have been desirable.
	Bad consultation experience	That you couldn't go back to a previous answer. You may think my previous choice would have been better, but then you can't go back into the comparison. I didn't like that in this consultation. Then put them in a row next to each other and it will be clearer.
	Too many/ little number of questions	Too many questions, questions are sometimes vague or too open, too much owning and having to formulate answers, too little explanation
	Too complex	Way too complicated. Too many options. Making a choice and then indicating whether it is yes or no is not easy. Dividing 100 points is also not clear and too complicated for many people.
	Only in Dutch	Still quite difficult to grasp and understand for many people. Idea: also offer in Frisian and English.
Transparency	Steering into certain direction	The questions are very guiding. The entire setting of the consultation already assumes that the line will be built anyway. Can't really be called objective.
	Others/ Trust Issue	that I have so little confidence in the government that I wonder whether this was useful....
	Security is questionable	Unlike a citizens' council, this is not yet sufficiently structured and formalized. So the results are not yet very reliable in my opinion. The subject can also be 'hijacked' by an interest group.
	Anonymous	Anonymous, there could be conscious group formation to push things in a certain direction.

I Broad Welfare Criteria

TNO and the Netherlands' Ministry of Infrastructure and Water Management (I&W) have developed building blocks of broad welfare, speficially utilized for evaluation of appraisal methods. The building blocks and explanation of each block is presented in Table I.1.

Table I.1: Broad Welfare Criteria for Assessment of Appraisal Methods

BW Criteria	Explanation	Question
Time periods: Taking history into account and adopting a more future-oriented perspective	What time period(s) are relevant? To what extent is it important to look at history or the future? What are the future impacts? For who and at what time horizon (next year or future generations)? How future-proof is the approach and the interventions?	Does the tool/approach make the time period explicit and allows for various time periods?
Area based	If the tool/approach highlights what geographical scales are important and what linkages between scales are relevant.	Does the tool/approach make the regions explicit and allows for various regions?
	What are the region-specific impacts? Do they differ between and within regions?	
Participation and co-creation	What is the desired level of participation? At what point is interaction appropriate? Does the method motivate for (higher levels of) participation and cocreation?	Does the tool/approach make the level of participation and co-creation explicit and allows for various levels of participation and co-creation?
Target groups and distribution effects	Which target groups feel the effects? Who has the benefits and burdens? How is this distributed within groups and between groups?	Does the tool/approach make the target groups and distribution effects explicit and provide insights on this?
Integral approach & systems perspective	Which topics are central? What is the connection between topics? What layers of government are involved?	Does the tool/approach make the breath explicit and allows for an (more) integral/ systems' perspective?
Synergies and trade-offs	Does the method motivate to create synergies or to make explicit trade-offs?	Does the method motivate to create synergies or to make explicit trade-offs?
	Are there opportunities for synergy (e.g. multiple use of space)? Are there trade-offs? Which topic gets priority and at the expense of which other topics?	