Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 97 (2026) 446–456

Transportation Research Procedia

www.elsevier.com/locate/procedia

13th International Conference on Transport Survey Methods

# Toward respondent validation of travel behavior segmentation results using a quantitative online self-assessment survey

Nicolas Salbach[a]*, Lukas Burger[a]*, Miriam Magdolen[a], Peter Vortisch[a]

[a]*Institute for Transport Studies, Karlsruhe Institute of Technology (KIT), Otto-Ammann-Platz 9, 76131 Karlsruhe, Germany*

**Abstract**

Segmentation is widely applied in quantitative travel behavior research to categorize heterogeneous populations into more homogeneous subgroups to create a more nuanced understanding of travel behavior and facilitate targeted policy and planning. However, without proper validation, identified segments may reflect artifacts of the data or the statistical method used rather than meaningful mobility patterns, limiting their practical use. Respondent validation, which assesses whether individuals recognize themselves in the segments they were assigned to, has received little attention in quantitative travel behavior research. This paper introduces a quantitative online self-assessment survey as a method to pursue respondent validation of segmentation results. Using a case study in Munich, Germany, we evaluate whether respondents can comprehend, differentiate, and identify with segment profiles, how well their self-assessments align with statistical assignments and membership probabilities, and what reasons they provide for mismatches. The results indicate that the profiles of segments created as part of our method were generally perceived as appealing, comprehensible, and distinguishable, though only a minority of respondents fully identified with a single segment. Exact matches of self-assessments and statistical assignments to segments were rare, but self-assessments correlated positively with membership probabilities, suggesting that probabilistic segmentations may be better suited for quantitative respondent validation than deterministic assignments. To reveal potentials for refinement of the segments and their descriptions, our method lastly enabled respondents to attribute discrepancies between self-assessments and statistical assignments to specific segment characteristics. Factors specific to our case study limit the generalizability of our findings. Overall though, the findings suggest that our method can provide systematic insights into respondents' evaluations of segmentation results and can thus contribute to the validation of segmentation results, especially strengthening their credibility, transparency, and practical relevance.



*Keywords:* quantitative respondent validation; segmentation; travel behavior; self-assessment; travel skeleton; urban mobility types

* Corresponding authors. *E-mail addresses:* nicolas.salbach@kit.edu, lukas.burger@kit.edu


10.1016/j.trpro.2026.04.182

## 1. Introduction and related work

### *1.1. Segmentation in quantitative travel behavior research*

Segmentation is a widely applied approach used across disciplines to categorize heterogeneous populations into more homogeneous subgroups based on shared attributes and/or behaviors. In travel behavior research, it is often used to facilitate a more nuanced understanding of transport demand in countries, regions, cities, or neighborhoods beyond aggregate averages. This is not only valuable for advancing theoretical perspectives on mobility but can also serve as an empirical foundation for targeted policies, demand management strategies, and tailored mobility services.

Segmentation studies in quantitative travel behavior research have drawn on a range of data sources. Building on these diverse data, various statistical segmentation methods have been applied. One of the most established methods is cluster analysis, which assigns each respondent to one segment (e.g., Anable, 2005; Lanzendorf, 2002). Over the last decade, probabilistic approaches like latent class cluster analysis have gained prominence (e.g., Molin et al., 2016; Rafiq and McNally, 2021). The resulting segments are often complex due to their multidimensionality, which poses challenges for comprehension, interpretation, and differentiation between segments, especially for non-researchers.

### *1.2. Validation of segmentation results in quantitative travel behavior research*

An important aspect, which only receives relatively little attention in many quantitative segmentation studies, is the validation of segmentation results. Without proper validation, however, identified segments may reflect artifacts of the data or the statistical method used rather than meaningful differences in travel behavior and/or other attributes, making them potentially unstable and of limited practical use.

The commonly reported form of quantitative validation is the *internal validation* of segmentation results, which assesses within-group homogeneity and between-group heterogeneity, ensuring that the segments capture meaningful patterns rather than random variation. Common approaches include cohesion and separation indices for cluster analyses and model-based criteria for latent class models. Some studies also examine the *reliability* of segmentation results by assessing whether individuals are consistently assigned to the same segments under repeated measurements, variations in the dataset, or alternative methodological specifications. *External validation* – testing the segments against other datasets or external outcomes – is conducted less frequently in quantitative travel behavior research, limiting the evidence for the generalizability and practical applicability of the segmentation results of many studies.

In addition to statistical examinations of validity, segmentation results are usually checked for their *clarity* and/or *interpretability*. However, such assessments typically rely solely on the researchers' subjective judgment. This can introduce researcher bias and does not ensure that the segments can be understood and interpreted by others – especially transport planners, policy makers, or the public – thus limiting their practical applicability.

### *1.3. Qualitative respondent validation of segmentation results*

In qualitative research, *respondent validation*, also referred to as *member checking*, is a commonly used method in which researchers present data or results back to respondents to check for accuracy and resonance with their own assessments. Given the potential influence of both respondent and researcher biases, member checks cannot be regarded as an unequivocal method of validation (Birt et al., 2016). Nevertheless, Lincoln and Guba (1985), whose work is widely considered as the gold standard for establishing trustworthiness in qualitative research, recommend member checking to minimize researcher bias. In an earlier work, Guba (1981) even describes member checks as the single most important action qualitative researchers can take to ensure the credibility of their analyses. Similarly, Stahl and King (2020) emphasize that respondent validation of any sort should lead toward trust in the research(ers).

In travel behavior research, several mixed-method studies have recontacted respondents following a quantitative segmentation to discuss their assigned segments. Rabe et al. (2002) and Gebhardt and Oostendorp (2021) conducted qualitative interviews with representatives of previously identified segments to characterize the segments in more detail and to gain additional insights into the underlying rationales of their travel behavior. However, these studies did not verify with respondents whether the segment assignments corresponded to their self-assessments. A rare example of a mixed-method study pursuing respondent validation of segments is the study of Hunecke and Haustein (2007), who conducted semi-structured interviews with representatives of previously identified segments to assess the fit.

### *1.4. Research objective and structure of the paper*

In contrast to qualitative methods, quantitative approaches require less time from both researchers and respondents, allow for larger samples, and enable statistical analyses, supporting the generalizability of the findings. To make use of these advantages, we present a method to collect respondents' self-assessments and opinions on the results of a prior quantitative segmentation through a quantitative follow-up online survey. Based on a case study in a neighborhood in Munich, Germany, we evaluate how our approach can contribute to the respondent validation of quantitative segmentation results. To this end, we examine three dimensions: (1) whether respondents can comprehend segments, distinguish between them, and identify with (exactly) one; (2) the extent to which respondents' self-assessments align with statistical membership probabilities and segment assignments; (3) the reasons for discrepancies between self-assessments and statistical assignments. To the best of our knowledge, this is the first attempt to address respondent validation of segmentation results in travel behavior research using a quantitative approach.

The paper is structured as follows: In Section 2 we give an introduction to our case study and data basis. Section 3 describes our method in detail. In section 4, we present and evaluate the results from the case study and discuss the use of our approach for respondent validation of segmentation results. Finally, Section 5 concludes the paper.

## 2. Data

### *2.1. Travel behavior survey in Munich's Dreimühlen quarter*

As a case study to develop, test and evaluate our method, we used the transformation project "Bestandsquartier der Zukunft – Dreimühlenviertel", which took place in Munich's Dreimühlen quarter in 2023. This project aimed to identify and test mobility measures for the neighborhood in collaboration with residents and local business owners.

As part of the project, a quantitative online travel behavior survey was conducted in fall 2023. Respondents were recruited via posters, social media posts, and a neighborhood chat group. The survey approach was based on the so-called *travel skeleton survey* concept by von Behren (2021). The survey asked respondents to report their travel behavior during a typical week and details of their most recent long-distance day trips and overnight trips. To account for determinants of vehicle ownership and mode use, it also included a section in which respondents provided self-assessments on mobility-related psychological statements drawn from the standardized items sets of Steg (2005) and Hunecke et al. (2021). Lastly, respondents' sociodemographic information was collected. In total, around 160 local residents voluntarily completed the survey, corresponding to 2.5% of the neighborhood's total population. However, the sample was not representative, as middle-aged, highly educated, and high-income adults were overrepresented.

### *2.2. Classification into the Urban Mobility Types*

The main purpose of the survey was to segment residents based on their mobility routines and attitudes, to better understand the travel demand in the neighborhood and derive suitable mobility transition measures. However, a multi-dimensional segmentation did not appear statistically feasible given the small sample size. Therefore, the respondents were classified into the segments from an existing segmentation based on data from a previous *travel skeleton survey* – the eleven so-called *Urban Mobility Types*. These were identified by Magdolen et al. (2019) using a two-step clustering approach (Ward method + k-means), based on a survey conducted in Berlin, San Francisco, and Shanghai between October 2016 and January 2017. The segmentation used three everyday travel indicators (*trip frequency*, *share of trips for work or education*, *share of trips by car*), one long-distance travel indicator *(estimated frequency of long-distance day trips and overnight trips with a minimum distance of 100 km*), and six travel-related psychological factors (*attitude toward public transport use in everyday life*, *car orientation, bicycle orientation*, *perceived obligation to use public transport*, *resilience with regard to privacy in public transit and adverse weather for cycling*, *and perceived need for everyday mobility*). For details on the cluster-forming variables, see Magdolen et al. (2019).

For the classification, a Random Forest model was trained on this data, achieving a test accuracy of 90%. Prior to classification, the Dreimühlen sample had to be reduced to 140 respondents due to missing values. For each of the respondents, the model determined membership probabilities for all eleven *Urban Mobility Types*, with the assigned type defined as the one with the highest membership probability.

## 3. Methodology

To collect respondents' self-assessments and opinions regarding the *Urban Mobility Types* and to compare them with the statistical assignments, we developed and administered a quantitative follow-up online survey. One month after the initial travel behavior survey had been concluded, respondents from the Dreimühlen sample were invited to participate via e-mail. The sole incentive provided was that respondents would learn their *Urban Mobility Type*.

The follow-up survey comprised three components: the respondent's self-assessment on the *Urban Mobility Types*, the comparison of self-assessment and statistically assigned type, and questions on reasons for a mismatch between the latter. This section describes the steps and design choices involved in developing the survey.

### *3.1. Segment description and visualization*

To ensure that the *Urban Mobility Types* could be understood by survey respondents as easily as possible and to keep the response burden low, our first step was to develop clear and appealing descriptions and visualizations of the types. As a basis, we used the values of the cluster-forming variables from the 1,662 respondents in the original clustering by Magdolen et al. (2019). Following Hunecke and Haustein (2007), we described each type only with the cluster-forming variables for which (almost) all representatives displayed a similar value (e.g., positive bicycle orientation). This approach aimed to reduce the complexity of the type descriptions by minimizing the number of descriptive dimensions, thereby also making the characteristics of each type more salient.

The data analysis revealed that the variable *car orientation* did not exhibit a distinct pattern for any of the types and that the psychological factor *perceived need for everyday mobility* was ambiguous in its interpretation. In addition, these two variables had the lowest variable importance score of all cluster-forming variables in the original clustering of Magdolen et al. (2019). Therefore, both variables were excluded from the descriptions. Overall, the analysis reduced the dimensionality of the type descriptions from originally ten attributes to between three and eight.

After dimensionality reduction, the remaining attributes were translated into standardized everyday language descriptions. For this, we analyzed the calculations underlying the four travel behavior indicators and interpreted the four remaining psychological factors based on the factor loadings reported by Magdolen et al. (2019). A challenge emerged with the *estimated frequency of long-distance day trips and overnight trips with a minimum distance of 100 km*. As it was partly based on non-mandatory questions, the exact values were deemed unreliable and only general tendencies were described. Finally, each attribute was visualized in standardized graphics in accordance with the descriptions. A final check ensured that each type remained distinctive through the dimensionality reduction, differing from all others in at least one characteristic. A subsequent pretest (n=12) revealed that graphical visualizations together with brief textual descriptions are the clearest and most appealing way to present the *Urban Mobility Types* to members of the general public. Figure 1 shows the resulting profiles for types 4, 5, and 9.

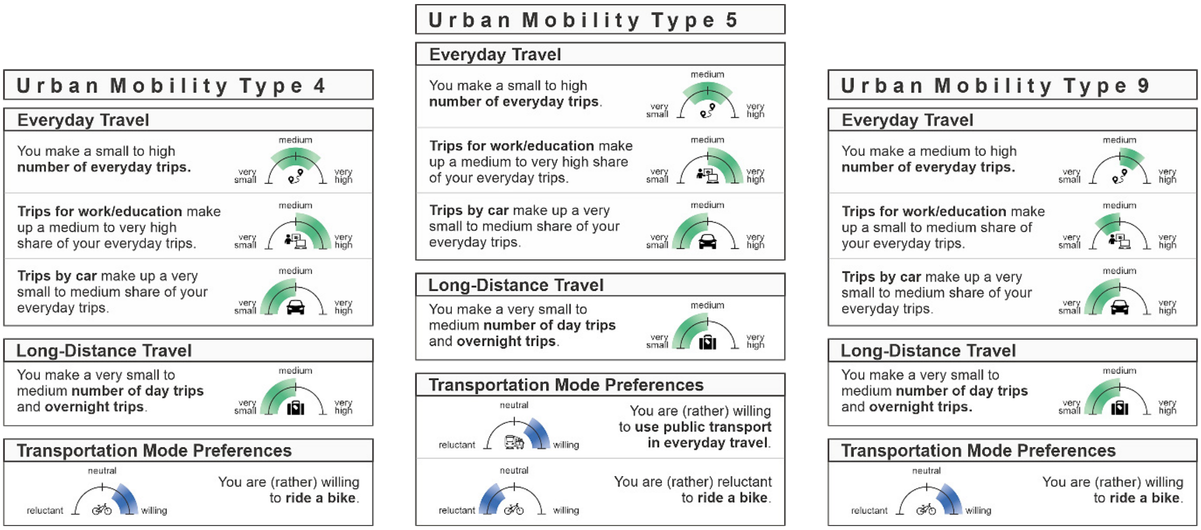


Figure 1. Selected Urban Mobility Type profiles

### *3.2. Rating method*

Our respondent validation approach aimed to assess whether respondents could evaluate their fit with the *Urban Mobility Type* profiles and distinguish between them. Since we anticipated it to be challenging to differentiate among eleven type profiles, the second step of our method was to select a suitable rating method for the self-assessments.

Therefore, in a second pretest (n=11), we assessed the suitability of two rating methods: a top-three ranking (three types with the highest fit in descending order) and an independent rating of the fit with each type on a five-point Likert scale (“fully inaccurate” to “fully accurate”). The pretest revealed a clear participant preference for the Likert scale due to the lower cognitive burden and the resulting shorter completion time. This preference also held when the number of types to be evaluated was reduced. Additionally, a Likert scale reveals the absolute identification with the types, a feature important for respondent validation, as it enables analyzing how strongly respondents identify with each type and whether they identify equally well with more than one type.

### *3.3. Survey procedure*

The second pretest also suggested that rating all eleven *Urban Mobility Types* would impose too high a response burden. To reduce this burden, we made only the self-assessment of six types mandatory and gave the option to voluntarily rate the remaining five. This required a personalized survey so that each respondent would (unknowingly) assess their assigned type within the six mandatory ratings.

At the start of the self-assessment, each respondent had to rate Type 1 (the only profile containing all eight attributes) to become familiar with the attribute set. Other than that, no explanation of the attributes was given. Respondents were then presented with five types in random order: the four with the highest individual membership probabilities and one random type. The respondents opting for the voluntary section were subsequently shown the remaining five types in random order. Information on which types to display and on the assigned type was transmitted to the survey software via a personalized URL for each respondent.

After completing the self-assessment, respondents were shown their assigned type and informed whether it matched their self-assessment. A match was declared when the assigned type was rated "fully accurate," regardless of ratings of other types. If self-assessment and assignment did not match, respondents were asked which characteristics did not align with their self-perception and to indicate the direction of misalignment ("too low" or "too high") or to provide another reason. At the end of the survey, all respondents could also provide important aspects of their mobility they felt were not captured by the *Urban Mobility Types*. Figure 2 illustrates the overall procedure of the follow-up survey.

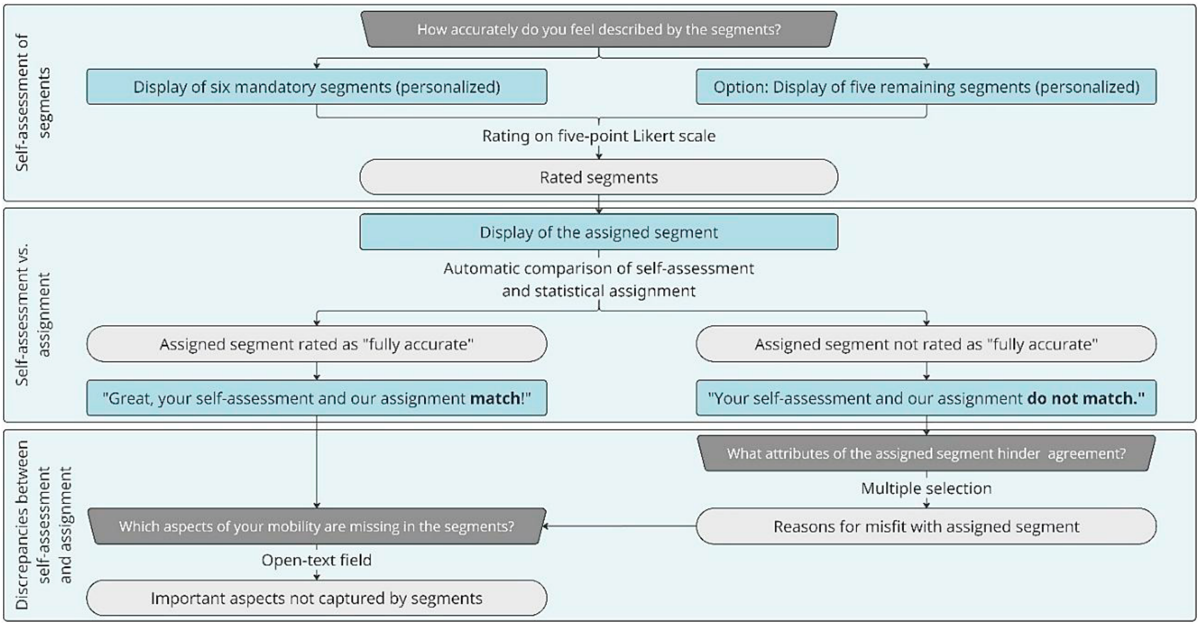



Figure 2. Procedure of the follow-up survey

## 4. Results and discussion

In the previous section, we presented our method to collect respondents' self-assessments and opinions on the results of a prior segmentation. In this section, we draw on analyses from our case study to evaluate how our approach can contribute to respondent validation of segmentation results. To this end, we examine three key aspects of respondent validation: (1) whether respondents can comprehend segments, distinguish them, and identify with (exactly) one; (2) the extent to which respondents' self-assessments align with statistical membership probabilities and segment assignments; (3) the reasons for discrepancies between self-assessments and statistical assignments.

### *4.1. Evaluation of respondents' self-assessments*

Of the 140 respondents from the Dreimühlen quarter assigned to the *Urban Mobility Types*, 116 consented to be re-contacted. In December 2023, they were e-mailed a personalized link to participate in the self-assessment survey. Of the 86 respondents who started the survey, 31 quit immediately upon reaching the self-assessment section and 53 completed it. Among these, 29 rated the mandatory six types, one respondent rated eight, and 23 voluntarily rated all eleven. The relatively high share voluntarily rating all eleven types suggests that the profiles were perceived as appealing and comprehensible by many respondents. However, a threefold self-selection bias (travel behavior survey participation, re-contact consent, and follow-up survey participation) limits the generalizability of this finding.

*Response times.* For the self-assessment section, response time was recorded and used to compute respondents' average response times (ART) per rated type. Table 1 shows the temporal statistics for 50 of the respondents. Three respondents, who exceeded 30 min for the completion, were treated as outliers and omitted from the temporal analysis.

Table 1. Statistics of average (mean) response time (ART) per type and respondent

| Number of rated types | Number of respondents | Mean of ART | Median of ART | Std. of ART |
|---|---|---|---|---|
| 6 | 27 | 35 s | 28 s | 18 s |
| 8 | 1 | 68 s | 68 s | - |
| 11 | 22 | 30 s | 28 s | 12 s |
| Total | 50 | 33 s | 28 s | 16 s |

The mean of the ART per type was 33 seconds, with a median of 28 seconds. Respondents who rated six types took five seconds longer per type on average than those who rated eleven. The minimum ART per type was 18 seconds among the six-type group and 11 seconds among the eleven-type group. Overall, the response times appear reasonable for comprehending the type profiles and providing self-assessments on them on a Likert scale. Shorter ART for the eleven-type group suggest a learning effect as respondents grew familiar with the structure and attributes of the profiles. However, recording the time spent on each rating would have allowed for more robust conclusions.

*Interpersonal analysis of self-assessments.* Figure 3 shows the distributions of respondents' self-assessments for all eleven *Urban Mobility Types*. Types 4 and 9 stand out with particularly positive evaluations, as around half of the respondents who assessed them rated them as "rather accurate" or "fully accurate." Both profiles are distinguished from the other types by having both a low share of car trips and a positive bicycle orientation. This aligns with findings from the travel behavior survey, which showed cycling (and walking) to be the most frequently used mode(s) and many respondents not owning a car. By contrast, Types 1 and 5 received predominantly negative evaluations, which is also in line with the travel behavior survey, as they are the only two types characterized by a negative bicycle orientation. The low acceptance of Type 1 is supported by the findings of Magdolen et al. (2019), who found this type almost exclusively in San Francisco. A similar pattern is evident for Type 10, whose low ratings likewise align with this earlier evidence. Type 8, characterized by a very high share of car trips combined with a negative attitude toward the use of public transport, was also mainly rated "rather inaccurate" or "fully inaccurate." Nevertheless, about one fifth of respondents evaluated this type positively, which aligns with the presence of a small subgroup of monomodal car users in the Dreimühlen sample, as indicated by the travel behavior survey.

In sum, the interpersonal analysis of self-assessments reveals marked differences in *Urban Mobility Type* ratings, which are plausible considering the findings of the travel behavior survey and Magdolen et al. (2019). Regarding our method, this suggests that the profiles enabled respondents to interpret and distinguish the types. This conclusion, though, is limited by the individualized survey design, which led to varying (numbers of) respondents rating the types.

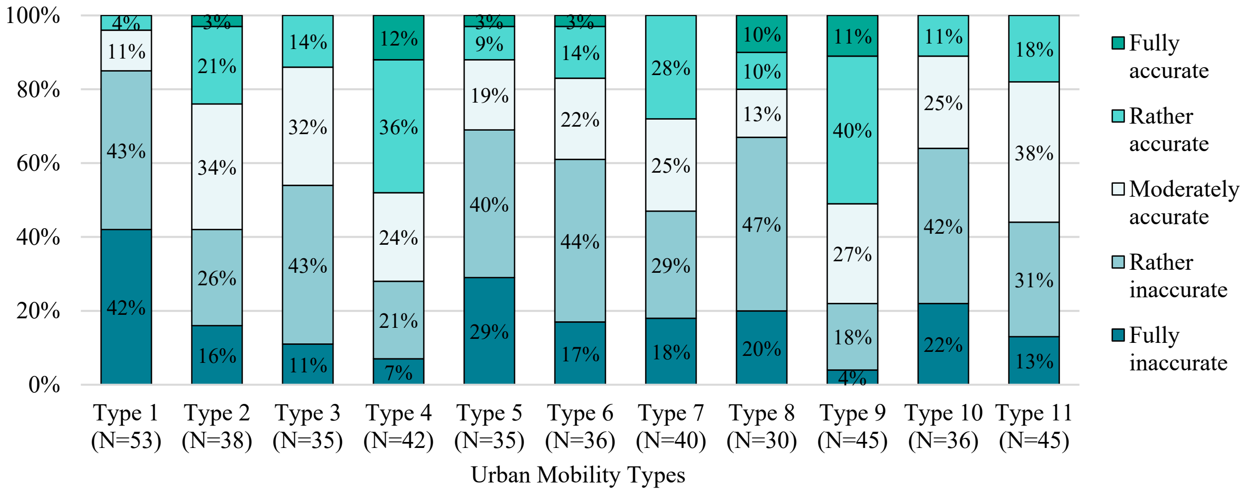



Figure 3. Distributions of respondents' self-assessments on the Urban Mobility Types

*Intrapersonal analysis of self-assessments.* A key aspect of respondent validation of the *Urban Mobility Types* is the evaluation of how well individuals can identify with (exactly one of) the types. Table 2 shows that of the 53 respondents only 12 fully identified with exactly one type, while 2 fully identified with two types. 17 respondents had a single preferred type, but without full alignment, and 11 felt that two types described them "rather accurately." An analysis of respondents with two top-rated types indicated that the profiles were often similar (e.g., Types 4 and 9). Ten respondents rated three or more types equally highest, most of whom had rated all eleven types. One respondent's highest rating was only "rather inaccurate," and three others reported a highest rating of "moderately accurate."

Table 2. Number of highest-rated types per respondent and corresponding highest rating(s)

| Number of highest-rated types of respondents | Highest rating(s) of respondents | | | | | Total |
|---|---|---|---|---|---|---|
| | Fully accurate | Rather accurate | Moderately accurate | Rather inaccurate | Fully inaccurate | |
| 1 | 12 | 17 | 0 | 0 | 0 | 29 |
| 2 | 2 | 11 | 0 | 1 | 0 | 14 |
| 3 | 0 | 6 | 2 | 0 | 0 | 8 |
| 4+ | 0 | 1 | 1 | 0 | 0 | 2 |
| Total | 14 | 35 | 3 | 1 | 0 | 53 |

Overall, the intrapersonal analysis of respondents' self-assessments reveals that more than half of the respondents gave the highest rating for a single type, and nearly four out of five rated at most two types highest. However, only few respondents could fully identify with a type. Thus, most respondents were able to distinguish between the *Urban Mobility Types*, but the majority of our sample did not feel fully represented by any one of them. However, according to the *Theory of Social Categorization* (Turner, 2010), self-assessment of individuals on group membership is often graded and overlapping rather than absolute, so the high proportion of "rather accurate" self-assessments as the highest rating cannot necessarily be interpreted as a bad fit of the sample to the types.

*Synthesis.* Our evaluation of respondents' self-assessments indicates that the *Urban Mobility Type* profiles were appealing, comprehensible, and distinguishable to most respondents. This can be concluded from the high willingness to voluntarily rate all profiles, reasonable response times, and general consistency of the self-assessments with the findings of the prior travel behavior survey and Magdolen et al. (2019). The profiles and the chosen rating approach proved successful as a basis for identifying that many respondents felt best represented by a single type, but only a few fully identified with one. This suggests that the *Urban Mobility Types* may not be fully suitable for describing mobility patterns of the Dreimühlen sample, though this interpretation is challenged by the *Theory of Social Categorization*. Overall, the findings suggest that the self-assessment survey can provide insights into respondents' comprehension of and identification with segments, indicating its usefulness for the respondent validation of segmentation results. However, generalizability is limited by the specifics of our case study, namely the small, selective sample and the individualized survey design.

### *4.2. Alignment of self-assessments with statistical segment assignments and membership probabilities*

To evaluate the extent to which respondents' self-assessments to the *Urban Mobility Types* aligned with the statistical assignments and membership probabilities, we treated the membership probabilities and type assignments generated by the Random Forest model as outcomes of a deterministic and probabilistic segmentation, respectively.

*Alignment of self-assessments and statistical assignments.* The first row of Table 3 shows the relationship between the statistically assigned type – defined as the type with the highest membership probability – and respondents' self-assessments of that type. Only for 3 of the 53 respondents a full match was attained, while another 16 rated their assigned type as "rather accurate." Thus, most self-assessments deviated (to varying degrees) from the statistical assignments. However, several factors specific to our case study may account for a major amount of this discrepancy: the independence of the samples, the determination of the *Urban Mobility Types* based on data from other cities, the seven-year gap between the two surveys (including the COVID-19 pandemic), and the Random Forest classification.

Table 3. Frequencies of statistical membership probabilities versus respondents' self-assessments of types

| Statistical membership probability of type | Ratings of respondents | | | | |
|---|---|---|---|---|---|
| | Fully accurate | Rather accurate | Moderately accurate | Rather inaccurate | Fully inaccurate |
| Highest | 3 (5.7%) | 16 (30.2%) | 17 (32.1%) | 13 (24.5%) | 4 (7.5%) |
| Second highest | 5 (9.4%) | 23 (43.4%) | 15 (28.3%) | 8 (15.1%) | 2 (3.8%) |
| Third highest | 2 (3.8%) | 10 (18.9%) | 14 (26.4%) | 19 (35.8%) | 8 (15.1%) |

| | | | | | |
|---|---|---|---|---|---|
| Other | 6 (2.2%) | 33 (12.0%) | 61 (22.1%) | 110 (39.9%) | 66 (23.9%) |

*Alignment of self-assessments and membership probabilities.* Table 3 further shows that self-assessments were more favorable for the types with the second-highest statistical membership probability than for the statistically assigned types, with more than half of the respondents rating them as "rather accurate" or "fully accurate." For the types with the third-highest membership probability, ratings were lower again and declined further for the remaining types. Although the direct match rate between statistical assignments and self-assessments was very low, this pattern suggests a positive association between membership probabilities and self-assessments.

To statistically assess this relationship, we used Spearman's rank correlation to compare membership probabilities and self-assessment ratings for each respondent. This method was selected to account for the metric membership probabilities, the ordinal Likert-scale ratings, and the occurrence of ties in the ratings. The results revealed a largely positive association: the mean of the correlation coefficients was 0.41, with a median of 0.50. Apart from three outliers with correlations below -0.6, all respondents showed positive values. A one-sided Wilcoxon signed-rank test against zero confirmed that the median correlation was significantly greater than zero ($V = 1249.5$, $p < 0.001$). In terms of respondent validation of the *Urban Mobility Types*, the analysis thus indicates that membership probabilities and respondents' self-assessments are generally aligned in their tendencies for many respondents, but not for all.

*Synthesis.* Respondents' self-assessments rarely matched their statistical type assignments, providing further evidence that the *Urban Mobility Types* may not be fully suitable for describing mobility patterns of the Dreimühlen sample. However, several factors specific to our case study likely contributed to the discrepancies. Most of these could be avoided by conducting the follow-up survey on the same sample the initial segmentation is based on, allowing for more robust conclusions on respondent validation of the segmentation results.

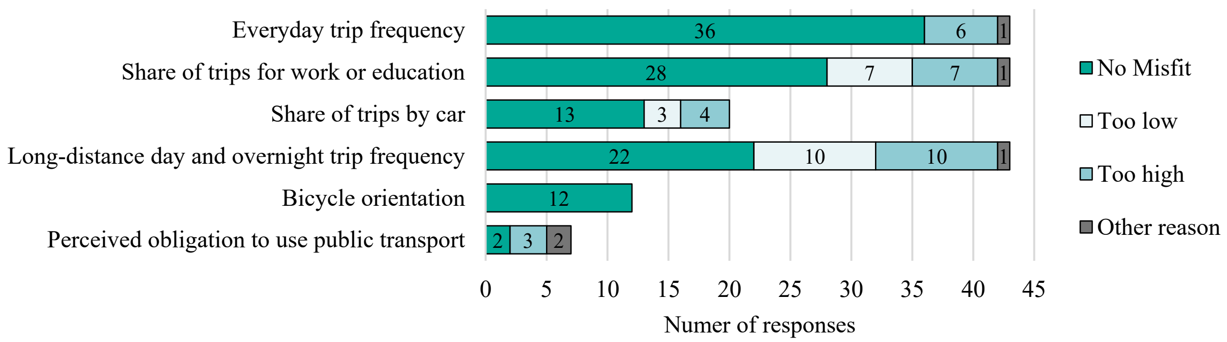
In contrast to the low match rate with type assignments, self-assessments showed a statistically significant positive rank correlation with the statistical membership probabilities. Considering the *Theory of Social Categorization*, this suggests that probabilistic segmentation results may be better suited for quantitative respondent validation than deterministic assignments, as they may better reflect the graded, overlapping nature of respondents' self-assessments toward different groups. However, due to our case study's specifics, this conclusion must also be drawn with caution.

### *4.3. Reasons for discrepancies between self-assessments and statistical assignments*

As the third aspect of respondent validation, we examined reasons for discrepancies between respondents' self-assessments and statistical assignments using the data from the final section of the follow-up survey. Of the 53 respondents who completed the self-assessment, seven quit after being shown their assigned *Urban Mobility Type*, and three had a match between their self-assessment and assigned type.

*Quantitative analysis of discrepancies by attributes.* Among the remaining 43 respondents with a mismatch, 32 identified a single attribute as the reason for the discrepancy, 7 selected two attributes, and 3 selected three attributes. One respondent stated in the open-text field that, upon reflection, they perceived no discrepancies with their assigned type, despite having rated it only as "rather accurate" in the self-assessment.

Figure 4 visualizes how frequently respondents indicated specific attributes as aligning or not aligning with their self-perceptions. The most common source of misalignment was the *estimated frequency of long-distance day trips and overnight trips of at least 100 km*, with 10 respondents each rating it as too low and too high. The *share of trips*



*for work or education* and the *share of trips by car* showed similar patterns, though with fewer mismatches. In contrast, all 12 respondents whose assigned type included the *bicycle orientation* agreed with this attribute, and 36 of 43 respondents concurred with the *trip frequency* in everyday travel. No assessments were provided for the attributes *perceived obligation to use public transport* and *resilience with regard to privacy in public transit and adverse weather for cycling*, as no respondent was assigned to a type including these attributes.

Figure 4. Respondents' reasons for discrepancies between self-assessments and statistical assignments

*Qualitative analysis of uncaptured aspects.* The analysis of 14 open-text responses on important mobility aspects respondents felt were not captured by the *Urban Mobility Types* yielded the following results: 6 respondents requested a more nuanced description of mode choice by distance, trip purpose, and/or season, 3 suggested including preferences between modes, and 3 took issue with the wording "very small share of trips by car", emphasizing they never use a car. These remarks highlight potentials for refining the *Urban Mobility Types* and their descriptions to make them more accessible, though any extensions of dimensionality entail the risk of adding complexity.

*Synthesis.* While only a minority of respondents fully identified with their assigned type, most mismatches were attributable to a single discrepant attribute. The *frequency of long-distance day trips and overnight trips of at least 100 km* was particularly contested, whereas *bicycle orientation* and everyday *trip frequency* were largely perceived accurate. The analysis also suggests that the possible interpretation based on the *Theory of Social Categorization* – that respondents who did not rate their assigned type as "fully accurate" might still have felt represented by it – does not hold, as all but one respondent with a statistical mismatch were able to pinpoint at least one attribute as the source of the discrepancy. However, given the case study specific uncertainties, these interpretations need to be made with caution. Regarding the general use of the last part of our survey for respondent validation, though, the analysis demonstrates that our approach can provide valuable systematic insights to refine segmentations and their descriptions.

## 5. Conclusion

This paper introduced a quantitative online self-assessment survey for the respondent validation of segmentation results in travel behavior research. Based on a case study in a Munich, we evaluated whether respondents could comprehend, differentiate, and identify with *Urban Mobility Type* profiles, how their self-assessments aligned with statistical assignments and membership probabilities, and what reasons they provided for discrepancies.

Our findings indicate that the profiles were generally appealing, comprehensible, and distinguishable, yet only a minority of respondents felt fully represented by a single type. Statistical analyses revealed a low rate of exact matches between type assignments and self-assessments, but a significant positive correlation of the latter with membership probabilities. This suggests that probabilistic segmentation results may be more suitable for quantitative respondent validation than deterministic assignments, as they better reflect the graded, overlapping nature of self-categorization. The discrepancy analysis showed that mismatches were often attributable to specific attributes rather than to entire profiles, highlighting concrete starting points for refinement of the segments and/or their descriptions. Importantly, almost all respondents with a statistical mismatch could identify at least one source of disagreement, supporting our decision to only declare a match with the assigned type if the corresponding self-assessment was "fully accurate."

Overall, our findings suggest that quantitative respondent validation of travel behavior segmentation results is feasible and useful. Our method can provide systematic insights into respondents' evaluations of segmentation results and can thus help validate or refine segments and their descriptions by highlighting contested attributes. At the same time, the conclusions from our case study must be interpreted with caution due to its limitations: the small, selective sample, the individualized survey design due to a high number of segments, and the separation of the survey from the original segmentation sample. Compared to qualitative approaches, another limitation of our method is the lower level of interpretive depth. While it can identify systematically contested segment assignments and attributes, it cannot provide definite causes for discrepancies. To uncover underlying reasons, a qualitative drill-down would be required.

Future applications of our method should be done with larger, more diverse samples, using the same data as the initial segmentation. Beyond travel behavior research, the method may also support respondent validation in other fields where segmentation is used to inform policy makers and/or practitioners by strengthening the credibility, transparency, and practical relevance of segmentation results.

## Author contributions

Nicolas Salbach: investigation, data curation, formal analysis, methodology, writing – original draft;
Lukas Burger: conceptualization, methodology, project administration, supervision, writing – review & editing;
Miriam Magdolen: validation, supervision; and Peter Vortisch: supervision

## References

Anable, J., 2005. 'Complacent Car Addicts' or 'Aspiring Environmentalists'? Identifying travel behaviour segments using attitude theory. Transp. Policy 12, 65–78. https://doi.org/10.1016/j.tranpol.2004.11.004

Birt, L., Scott, S., Cavers, D., Campbell, C., Walter, F., 2016. Member Checking: A Tool to Enhance Trustworthiness or Merely a Nod to Validation? Qual. Health Res. 26, 1802–1811. https://doi.org/10.1177/1049732316654870

Gebhardt, L., Oostendorp, R., 2021. Alles eine Frage der Logik?! Erkenntnisse einer Mixed-Method-Studie zur Pkw-Nutzung in Berlin. Geogr. Helvetica 76, 115–127. https://doi.org/10.5194/gh-76-115-2021

Guba, E.G., 1981. Criteria for assessing the trustworthiness of naturalistic inquiries. ECTJ 29, 75–91. https://doi.org/10.1007/BF02766777

Hunecke, M., Haustein, S., 2007. Einstellungsbasierte Mobilitätstypen: Eine integrierte Anwendung von multivariaten und inhaltsanalytischen Methoden der empirischen Sozialforschung zur Identifikation von Zielgruppen für eine nachhaltige Mobilität. Umweltpsychologie 11, 38–68.

Hunecke, M., Heppner, H., Groth, S., 2021. Fragebogen zu psychologischen Einflussfaktoren der Nutzung von Pkw, ÖPNV und Fahrrad (PsyVKN): Faktorenstruktur, psychometrische Eigenschaften und Validierung. Diagnostica 68, 3–13. https://doi.org/10.1026/0012-1924/a000277

Lanzendorf, M., 2002. Mobility Styles and Travel Behavior: Application of a Lifestyle Approach to Leisure Travel. Transp. Res. Rec. 1807, 163–173. https://doi.org/10.3141/1807-20

Lincoln, Y.S., Guba, E.G., 1985. Naturalistic Inquiry. SAGE.

Magdolen, M., von Behren, S., Chlond, B., Hunecke, M., Vortisch, P., 2019. Combining Attitudes And Travel Behavior: A comparison of urban mobility types identified in Shanghai, Berlin and San Francisco, in: 98th Annual Meeting of the Transportation Research Board. https://doi.org/10.5445/IR/1000093783

Molin, E., Mokhtarian, P., Kroesen, M., 2016. Multimodal travel groups and attitudes: A latent class cluster analysis of Dutch travelers. Transp. Res. Part Policy Pract. 83, 14–29. https://doi.org/10.1016/j.tra.2015.11.001

Rabe, S., Bäumer, D., Tully, C.J., Hunecke, M., 2002. Das Projektdesign von U.Move, in: Hunecke, M., Tully, C.J., Bäumer, D. (Eds.), Mobilität von Jugendlichen: Psychologische, soziologische und umweltbezogene Ergebnisse und Gestaltungsempfehlungen. VS Verlag für Sozialwissenschaften, Wiesbaden, pp. 39–46. https://doi.org/10.1007/978-3-663-01176-7_2

Rafiq, R., McNally, M.G., 2021. Heterogeneity in Activity-travel Patterns of Public Transit Users: An Application of Latent Class Analysis. Transp. Res. Part Policy Pract. 152, 1–18. https://doi.org/10.1016/j.tra.2021.07.011

Stahl, N.A., King, J.R., 2020. Expanding Approaches for Research: Understanding and Using Trustworthiness in Qualitative Research. J. Dev. Educ. 44, 26–28.

Steg, L., 2005. Car use: lust and must. Instrumental, symbolic and affective motives for car use. Transp. Res. Part Policy Pract., Positive Utility of Travel 39, 147–162. https://doi.org/10.1016/j.tra.2004.07.001

Turner, J.C., 2010. Social categorization and the self-concept: A social cognitive theory of group behavior, Rediscovering social identity. Psychology Press, New York, NY, US.

von Behren, S., 2021. Das Mobilitätsskelett – ein integrativer Ansatz zur mehrdimensionalen Betrachtung von urbaner Mobilität (Dissertation). Karlsruher Institut für Technologie, Karlsruhe.