



Strategic Roads User Survey

Analysis of key drivers of overall satisfaction

August 2020

This report was written by Jana Kühne and Gill Edwards at Kantar Analytics Practise for Transport Focus



Contents

Summary	3
Key drivers of satisfaction	4
1.1 The approach	
1.2 Interpreting Gamma	
1.3 Data preparation and checks	
1.4 Results	
Appendix	7

Summary

The objective of this analysis was to determine the key drivers of overall satisfaction using two years of SRUS survey results being available. The initial key driver analysis had been conducted based on the first 7 months of data (April to Nov 2018).

The analysis was conducted using a 'Gamma test' technique. This is a nonparametric bivariate test which examined the strength of relationship between the 17 most appropriate independent variables and overall satisfaction (further detail about methodology is in the Key Drivers of Satisfaction section of the report).

The table below shows the top five drivers and their strength of relationship with overall satisfaction (these terms being explained later in the report). For reference, the rankings in the initial analysis are shown for comparison.

Variable and question number in 2018-2020 questionnaire	Rank from 2018-2020 data (1 most associative)	Relationship with overall satisfaction	Rank in initial analysis
Satisfaction with journey time (Q41).	1	Strong	1
Whether other delays experienced (Q51).	2	Good	4
Actual journey time compared to expectation (Q40).	3	Good	2
Level of traffic on selected road (Q42).	4	Good	3
Satisfaction with roadworks management (Q49).	5	Good	6

Satisfaction with journey time is confirmed as having the strongest relationship with overall satisfaction, and the second to fifth rank exhibit a good relationship. While the results are not directly comparable with the initial analysis due to a change in methodology, they show the top five drivers of overall satisfaction are broadly consistent.

1. Undertaking the analysis and results

1.1 The approach

The early 2019 analysis was run by calculating a Chi-Square statistic for various measures against an overall satisfaction measure to establish the key drivers of satisfaction. It was run on April – November 2018 weighted data. Extensive exploratory analysis was carried out to identify questions influencing overall satisfaction from all questions asked within SRUS. The questions deemed relevant were selected; they involved a mix of nominal and ordinal response types with varying levels of missing data due to routings in the questionnaire. Chi-Square was considered the most appropriate test to apply to this mix of variables, being a nonparametric bivariate test. While that test approach provided a rank order of questions, it inherently cannot quantify the contribution of each variable to overall satisfaction. To address this aspect a subsequent step, a linear regression model was created to quantify the combined contribution of all statements to variability in overall satisfaction.

Deciding the approach to take in 2020 took account of the fact that the final 17 most relevant input variables used in the 2019 analysis were ordinal or dichotomous. It was therefore felt that a Gamma test should be calculated for each pair of variables in place of the previous Chi-Square approach and linear regression modelling. Gamma, also a nonparametric bivariate test, is a measure of the strength of the relationship between ordinal variables. It is widely used, easy to interpret as its outputs resemble the more commonly known correlation range of values and it provides the directionality of the relationship.

Gamma provides some key enhancements on the previous approach:

- Response scales do not require transformation and can be left in their original format, allowing for as much granularity as possible. In contrast, for results to be comparable across tests, Chi-Square analysis requires variables to be transformed into binary form which necessitated a reduction in the gradation within the information.
- The Gamma test approach within itself provides a quantification of the strength of relationship of each individual question with overall satisfaction; neither Chi-Square nor the statistical modelling allowed this.

1.2 Interpreting Gamma

Results from the Gamma analysis can be interpreted in the same way as a correlation with values that range from -1.00 to 1.00. A Gamma of 0.00 reflects no association; a Gamma of 1.00 reflects a perfect positive relationship between variables; a Gamma of -1.00 reflects a perfect negative relationship between those variables. As a rule of thumb, based on the Gamma values obtained one can classify the strength of the relationship of each potential 'driver-variable' into weak, good and strong (these being either positive or negative values as shown below):

- 0 and 0.39 – weak relationship (red shading)
- 0.4 and 0.69 – good relationship (yellow shading)
- 0.7 and 1 – strong relationship (green shading)

1.3 Data preparation and checks

In 2019 exploratory analysis was run to:

- establish potential drivers of satisfaction from the questionnaire. The variables in scope were based on their potential to provide actionable insights. Variables excluded were factors such as demographics, frequency of usage and journey purpose – even though they might be contributing to overall satisfaction.
- audit the data to identify variables whose values were usable: questions where respondents had almost universally answered either positively or negatively, or universally said either yes or no, were not considered as they would not be differentiating. An example is of this was “How safe felt travelling” where 92.9% of respondents had responded agree/agree strongly. However, ‘whether saw poor driving by other road users’ was included in the analysis; this variable’s values being more differentiating.
- run correlations of the usable variables remaining with the dependent variable. Independent variables with correlations less than a +/- 0.1 threshold were excluded from further analysis due to the weak relationship and so unlikely to impact overall satisfaction.

That process produced a list of 17 variables to bring into the assessment. The same variables were chosen for the full analysis of the latest available dataset, comprising journeys made from April 2018 to mid-March 2020. All the variables were handled in the same way as previously, with don’t know and can’t remember responses imputed with mid-points as appropriate (see Appendix 1 for further detail on the imputations made). The data was recoded to range from low to high, with overall satisfaction ranging from 1 (very dissatisfied) to 5 (very satisfied).

As a further quality check of moving to a Gamma model, the 2019 Chi-Square approach was re-run on the two years of data. That provided a ranking closely in line with results presented last year and indicating an acceptable consistency between the approaches.

1.4 Results

The table below shows the Gamma coefficient of the strength of association between individual variables and the overall satisfaction variable. They are ranked on absolute impact, from most impactful to least impactful.

Satisfaction with journey time has the strongest relationship with overall satisfaction. This is a positive association: those more satisfied with journey time are more likely to be satisfied overall.

The second to fourth ranks exhibit a ‘good’ inverse relationship with overall satisfaction. This can be interpreted as follows: if other delays are experienced / the longer the actual journey time compared to expectation / the higher the level of traffic, then overall journey satisfaction is lower.

The last item in the top five ranked variables is satisfaction with roadworks management which has a good positive relationship with overall satisfaction. However, it is worth noting that passing through roadworks has only a weak relationship with overall satisfaction.

Variable and question number in 2018-2020 questionnaire	Rank from 2018-2020 data (1 most associative)	Strength of relationship	Gamma	Values	Sample size
Satisfaction with journey time (Q41).	1	Strong	0.82	1 – very dissatisfied to 5 – very satisfied	16762
Whether other delays experienced (Q51).	2	Good	-0.66	1 – No 2 – Yes	16761
Actual journey time compared to expectation (Q40).	3	Good	-0.64	1 – less to 3 – more	16760
Level of traffic on selected road (Q42).	4	Good	-0.59	1 – light to 4 – congested	16714
Satisfaction with roadworks management (Q49).	5	Good	0.57	1 – very dissatisfied to 5 – very satisfied	4094
Satisfaction with information about roadworks (Q48).	6	Good	0.51	1 – very dissatisfied to 5 – very satisfied	4116
Satisfaction with the way delays were managed (Q57).	7	Good	0.51	1 – very dissatisfied to 5 – very satisfied	1973
Satisfaction with road surface (Q63 part 1).	8	Good	0.50	1 – very dissatisfied to 5 – very satisfied	16761
Satisfaction with information provided on electronic message signs (Q65).	9	Good	0.48	1 – very dissatisfied to 5 – very satisfied	10801
Satisfaction with information on permanent signs (Q64).	10	Good	0.44	1 – very dissatisfied to 5 – very satisfied	16759
Satisfaction with the information provided about delays (Q56).	11	Good	0.42	1 – very dissatisfied to 5 – very satisfied	1896
How satisfied are you with “Cat’s eyes” (Q63 part 4).	12	Good	0.42	1 – very dissatisfied to 5 – very satisfied	2555
Whether passed through roadworks (Q43).	13	Weak	-0.37	1 – No 2 – Yes	16759
How satisfied are you with “Road lighting” (Q63 part 2).	14	Weak	0.35	1 – very dissatisfied to 5 – very satisfied	2857
Whether saw poor driving by others (Q59 part 1).	15	Weak	-0.29	1 – No 2 – Yes	16760
Satisfaction with“Road and verge are free of litter” (Q63 part 5).	16	Weak	0.28	1 – very dissatisfied to 5 – very satisfied	16762
Agreement to statement “I enjoy driving” Q69.	17	Weak	0.19	1 – strongly disagree to 5 – strongly agree	16760

Appendix

Appendix - Imputation of missing values

Don't know and can't remember responses were replaced with mid-points where appropriate and for one variable (passed through road works) it was recoded as "no". Generally, don't know and can't remember responses represented only a small proportion, as can be seen below. The only exception was "satisfaction with the absence of litter" where just under a fifth of those asked recorded a don't know or can't remember response. A mid-point was still used to replace these responses to keep consistency of approach across all the variables affected. Variables treated in this way were:

- Satisfaction with journey time: "Don't know/can't remember" (0.2%) recoded as "Neither satisfied nor dissatisfied".
- Expected time journey took: "I had no expectation of journey length" (2.1%) and "Don't know/can't remember" (0.4%) recoded as "The same time as expected"
- Satisfaction with road surface: "Don't know/can't remember" (1.2%) recoded as "Neither satisfied nor dissatisfied"
- Satisfaction with road and verge being free of litter: "Don't know/can't remember" (19.6%) recoded as "Neither satisfied nor dissatisfied".
- Satisfaction with information provided on permanent signs: "Don't know/can't remember" (2.0%) and "Did not see such signs" (4.1%) recoded as "Neither satisfied nor dissatisfied"
- I enjoy driving on motorways: "Don't know/can't remember" (0.3%) recoded as "Neither agree nor disagree"
- Passed through any roadworks on journey: "Don't know" (5.4%) recoded as "No"