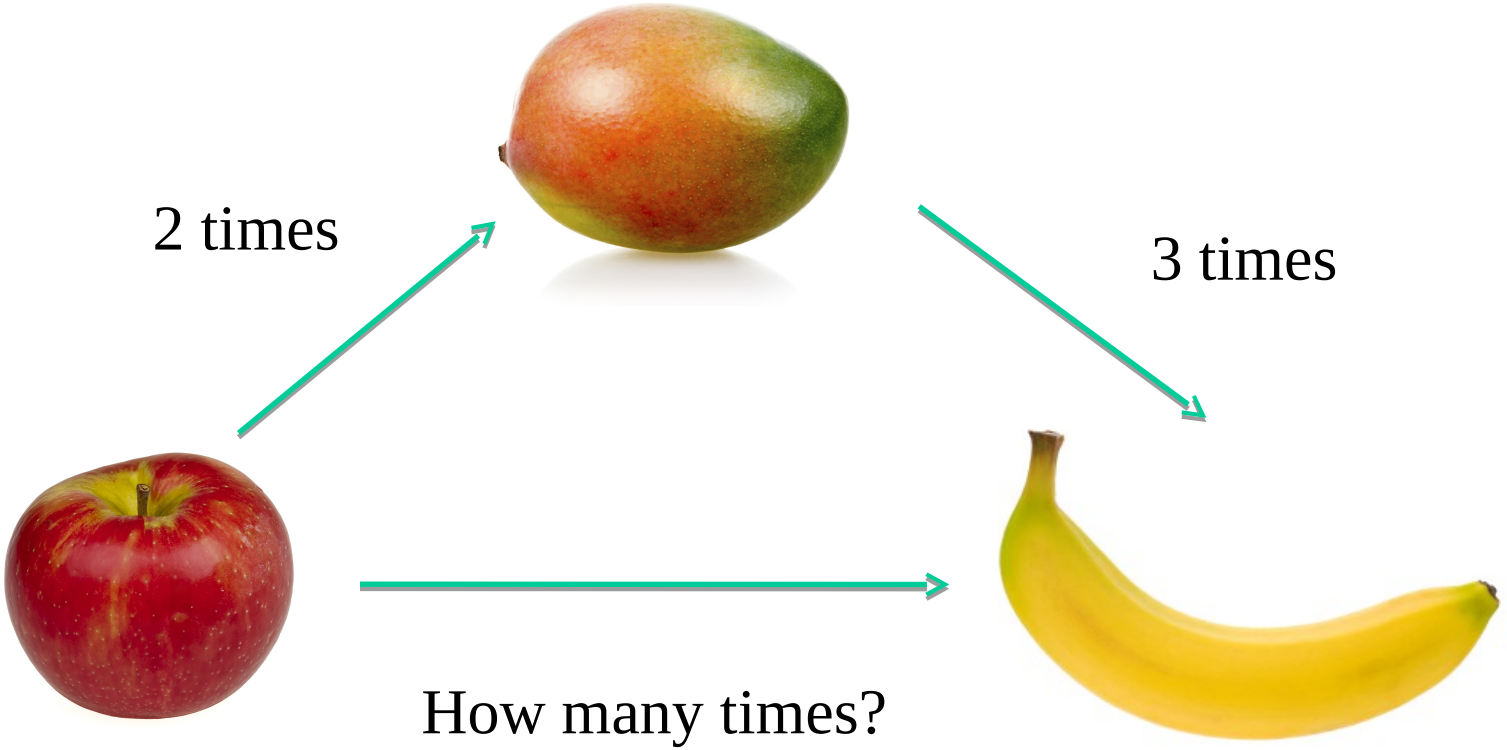
Are inconsistent decision better? An interactive experiment with pairwise comparisons

Portsmouth Business School
Richmond Building
Portland Street
PO1 3DE

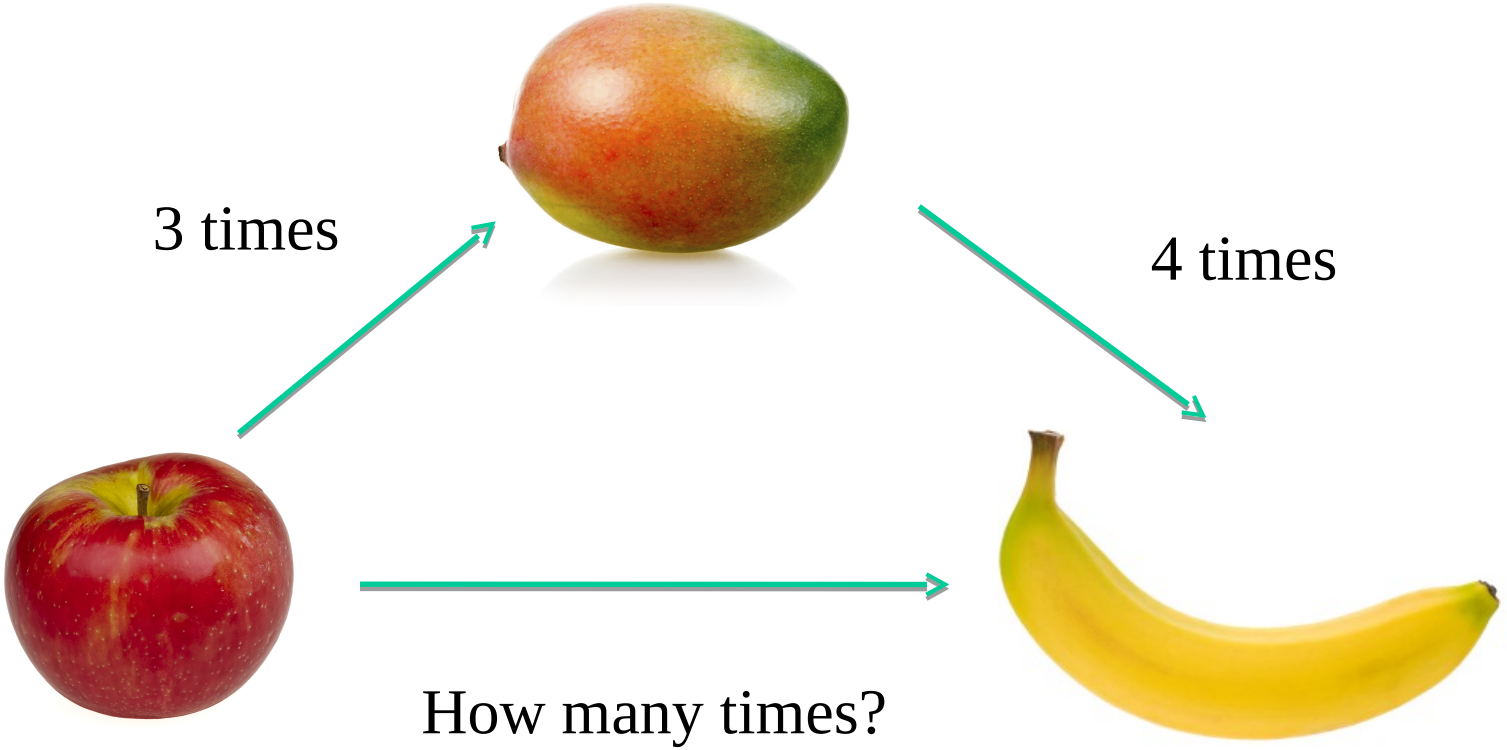
Alessio Ishizaka

Kyriacos Antoniades

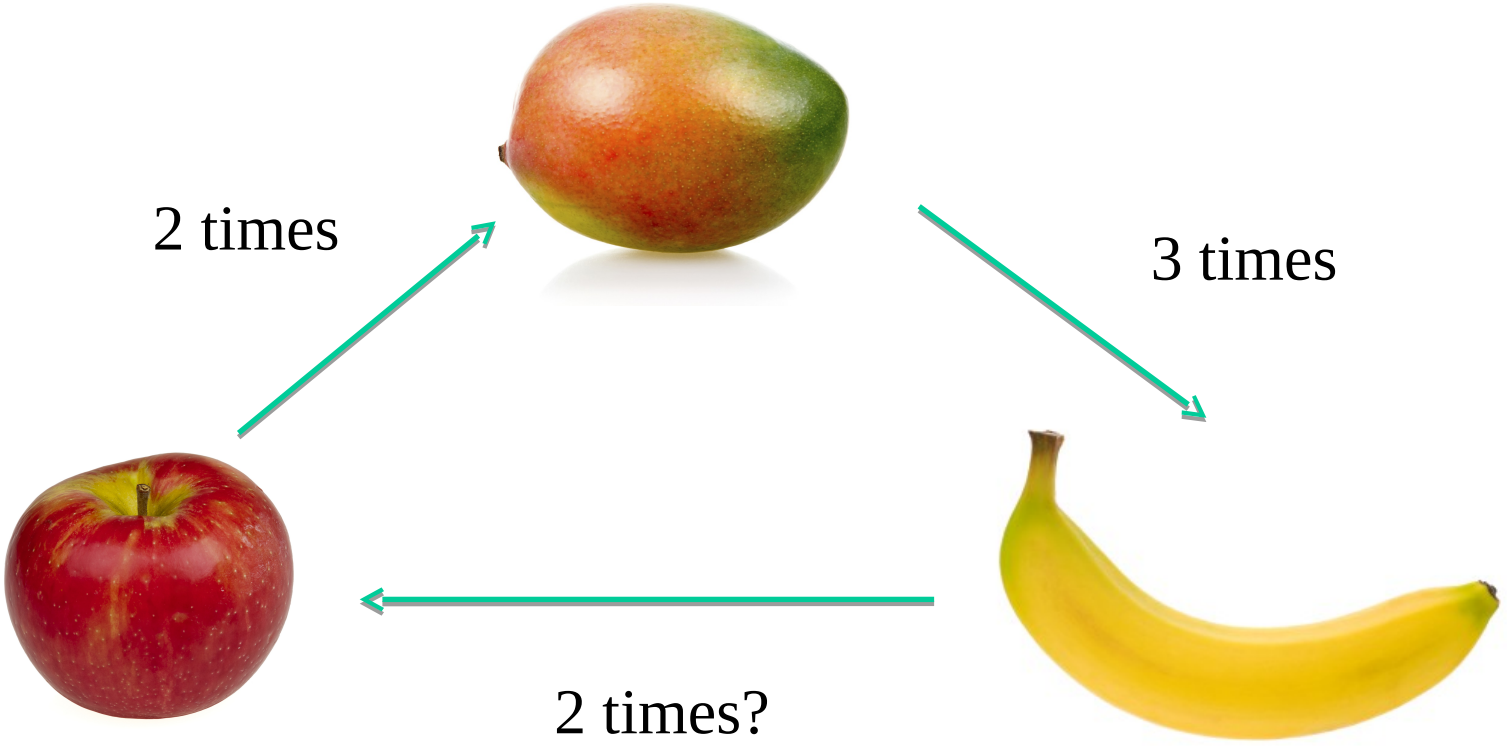
Which fruit do you prefer?



Which fruit do you prefer?



Is this inconsistency possible?



Comparison matrix

1	a_{12}		
a_{21}	1		a_{ij}
		1	
	$1/a_{ij}$		1

$$\mathrm{CI} = \frac{\lambda_{\max} - n}{n - 1}$$

where λ_{max} is the maximal eigenvalue

n is the dimension of the matrix

Consistency Ratio

- CR = CI/RI < 10%

where CR is the consistency ratio

RI is the random index

Saaty (1977) calculated the following random indices:

n	3	4	5	6	7	8	9	10
RI	0.58	0.9	1.12	1.24	1.32	1.41	1.45	1.49

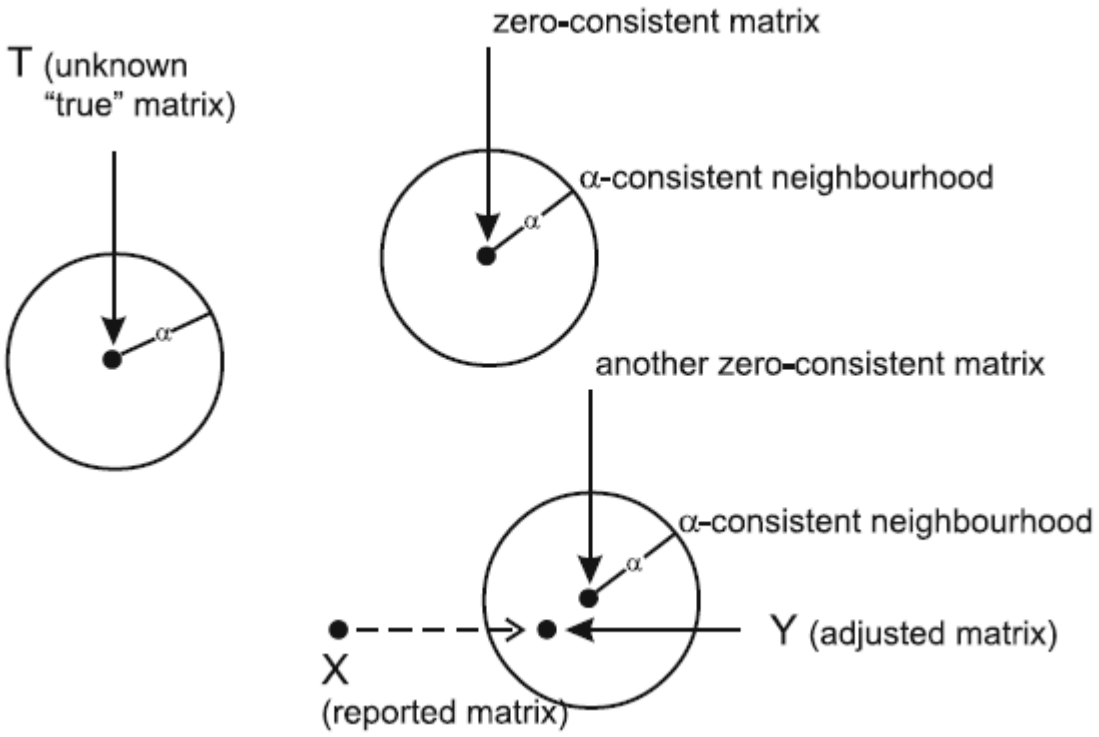
Critic: 10% is an arbitrary value

Inconsistent matrices are not better!

- 18 graduate students compare five different compact cars in global terms, and also in terms of their aesthetics
- When intransitivities are automatically removed, the preferences of decision makers are not better represented

- Linares, P. (2009). Are inconsistent decisions better? An experiment with pairwise comparisons. European Journal of Operational Research, 193(2), 492-498.

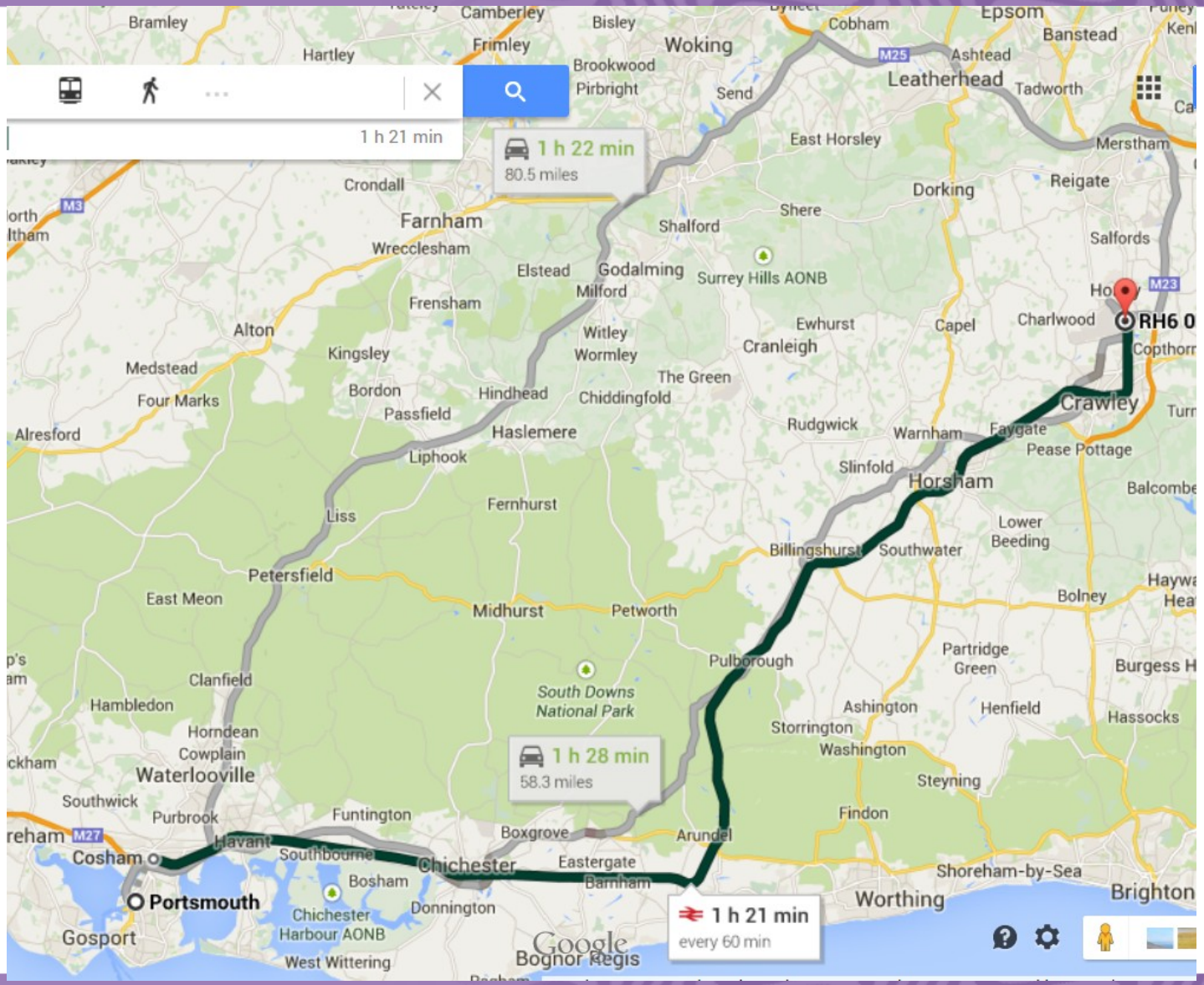
Why automatic correction does not work?



Gaul, W., & Gastes, D. (2012). A note on consistency improvements of AHP paired comparison data. Advances in Data Analysis and Classification, 6(4), 289-302.


Kirkwall
Thurso
Wick
Stornoway
Elgin
Fraserburgh
Inverness
Nairn
Peterhead
Kirkcaldy
Aberdeen
ATLANTIC OCEAN
Fort William
Scotland
Forfar
Arbroath
North Sea
Perth
Dundee
Oban
Stirling
Dunfermline
Glasgow
Edinburgh
Kilmarnock
Douglas
Ayr
Cumnock
Hawick
Londonderry
Coleraine
Dumfries
Newcastle upon Tyne
Carlisle
Stranraer
Ballymena
Bangor
Sunderland
Durham
Omagh
Belfast
Workington
Middlesbrough
Darlington
Isle of Man
Douglas
Scarborough
Bridlington
Newry
Barrow-in-Furness
England
York
Northern Ireland
Bradford
Kingston upon Hull
Irish Sea
Leeds
Preston
Manchester
Grimsby
Liverpool
Holyhead
Lincoln
Sheffield
Stoke-on-Trent
Nottingham
Derby
Leicester
Norwich
Wolverhampton
Great Yarmouth
Peterborough
Aberystwyth
Coventry
Cambridge
Birmingham
WEST MIDLANDS
Northampton
Luton
Colchester
Ipswich
Wales
Oxford
Milford Haven
Newport
London
Southend-on-Sea
Reading
Swansea
Margate
Cardiff
Bristol
Maidstone
Ashford
Taunton
Southampton
Brighton
Celtic Sea
Exeter
Bournemouth
Portsmouth
Weymouth
Plymouth
Torquay
Penzance
Truro
English Channel
0
200 km

University of Portsmouth



How to travel to Gatwick airport?

coach

train

personal car

car sharing

taxi

Rankings produced for each participant

- ***Original Ranking*** (R_O), where the priorities are calculated by the eigenvector method without any inconsistency correction.
- ***Automatic Ranking*** (R_A), where inconsistencies are corrected automatically using the goal programming method.
- ***Interactive Ranking*** (R_I), where the software indicates to the participant the most inconsistent pairwise comparison to the least one and invite her/him to change them.

Experimental procedure (1)

1. The decision problem is explained to the participant.
2. The participant pairwise compares the five alternatives.
3. The **Consistency Ratio**, the **Original Ranking** and the **Automatic Ranking** are calculated.
4. If the consistency ratio is acceptable, i.e. **below 10%**, the experiment **terminates** otherwise the consistency error ε_{ij} of each pairwise comparison, is calculated with (Saaty, 2003):

$$\varepsilon_{ij} = \max\left(a_{ij} * \frac{p_j}{p_i}, a_{ji} * \frac{p_i}{p_j}\right)$$

Experimental procedure (2)

5. **Possibility to revise** the most inconsistent comparison, i.e. the comparison with the highest ε_{ij}.

   If they decline, they are asked if they want to revise the next most inconsistent comparison.

   When they revise, the process restarts from point 4 until the inconsistency falls below 10% or the participant has considered all entries.

6. The final **Interactive Ranking** is calculated.

7. The participant is asked which of the three rankings (without knowing how they have been calculated) represents their preference.

Results: problem order influence

- Sixty-two participants.
  - The first thirty-one participants solved the problem with the subjective criterion first and then the problem with the objective criterion.
  - The next thirty-one participants solved the problems in the reverse order.
- Both samples produced statistically identical outcomes.
  - Order did not have any influence on the results

Consistency improvement with the interactive method

	Final matrix with improved consistency	Final matrix consistency improved not improved
Problem with subjective criterion	39	0
Problem with objective criterion	34	0

In 100% of the cases, the interactive method improved the consistency.

Alternatives ranking

City	Original Ranking	Interactive Ranking	Automatic Ranking	Normalised true distance
Cardiff	0.186 ± 0.069	0.190 ± 0.069	0.226 ± 0.081	0.162
London	0.260 ± 0.053	0.261 ± 0.053	0.267 ± 0.058	0.267
Edinburgh	0.046 ± 0.040	0.046 ± 0.041	0.068 ± 0.056	0.022
Southampton	0.423 ± 0.104	0.419 ± 0.093	0.338 ± 0.087	0.471
Liverpool	0.085 ± 0.033	0.084 ± 0.032	0.101 ± 0.036	0.078

- All three rankings ordered the distance of the cities to Portsmouth correctly.

Error between estimated and true distances

City	|Original priority -true distance|	|Interactive priority - true distance|	|Automatic priority - true distance|
Cardiff	0.087	0.082	0.140
London	0.057	0.058	0.089
Edinburgh	0.042	0.040	0.048
Southampton	0.024	0.024	0.046
Liverpool	0.023	0.022	0.029

- The automatic ranking is furthest from the true distance.
- The original and interactive priorities are very close, which makes the effort to improve the consistency questionable, if the final result does not improve.

Priorities of transport selection

City	Original priority	Interactive priority	Automatic priority
Train	0.187 ± 0.131	0.190 ± 0.132	0.183 ± 0.123
Coach	0.121 ± 0.104	0.127 ± 0.099	0.131 ± 0.098
Taxi	0.179 ± 0.142	0.173 ± 0.145	0.167 ± 0.133
Car sharing	0.176 ± 0.134	0.181 ± 0.137	0.161 ± 0.105
Own car	0.337 ± 0.219	0.214 ± 0.230	0.357 ± 0.229

- The priorities of the subjective problems were more dispersed, i.e. the standard deviation was higher .
- The “own car” alternative was by far the most preferred transportation mode in the original and interactive ranking. This clear preference for ‘own car’ was faded in the interactive ranking.

Participants' preferred ranking for the subjective problem

	Participants' preference of rankings		
	Interactive	Automatic	Original
Observed Frequency	13	14	12
Expected Frequency (proportion)	13 (.3̄)	13 (.3̄)	13 (.3̄)

Note. $\chi 2 = 0.15$, degree of freedom =2, significance threshold $p > .05$

A Chi-square test confirms that the frequencies of participants' preferences were equally distributed

Participants' preferred ranking for the objective problem

	Participants' preference of rankings		
	Interactive	Automatic	Original
Observed Frequency	18	4	12
Expected Frequency (proportion)	11.3 (.3̄)	11.3 (.3̄)	11.3 (.3̄)

Note. $\chi 2 = 8.71$, degree of freedom =2, significance threshold $p > .05$

- A Chi-square test confirmed that the frequency of the participants' preferences were not equally distributed.
- If the automatic ranking is ignored, there is no significant difference between the original and interactive ranking with a Chi-square test.

Closest ranking to the true value

	Rankings closest to the true value		
	Interactive	Automatic	Original
Observed Frequency	6	7	18
Expected Frequency (proportion)	10.3̄ (.3)	10.3̄ (.3)	10.3̄ (.3)

Note. $\chi 2 = 8.71$, degree of freedom =2, significance threshold $p > .05$

- A Chi-square test confirmed that the frequency of the participants' preferences were not equally distributed.
- If the automatic ranking is ignored, there is no significant difference between the original and interactive ranking with a Chi-square test.

Conclusions:

- **Consistency improvement:**
  - **The interactive and automatic methods *improved* consistencies in pairwise comparisons**
  - **Revisions *are in agreement* with the best fit for the pairwise comparison**
- **Representation of ranking :**
  - **Interactive approach *does not* better represent participants' preferences**
  - **The original ranking is closest to the true value in the objective problem**
  - **The difference between the priorities of the original and interactive ranking were found to be very small**

Main Conclusion

The effort to reduce inconsistencies using the interactive approach is questionable.

Dr Alessio Ishizaka
Reader in Decision Analysis
University of Portsmouth
Portmsouth Business School
Richmond Building
PO1 3DE Portsmouth
United Kingdom

Alessio.Ishizaka@port.ac.uk

Thank you!