


Using Statistical Methods to Access the Student Voice

Dr David Carroll

Manager, Strategic Analysis

Portfolio of the DVC (Education)

So, you have your data...

- Descriptive statistics (mean, median, % satisfied, etc.)
 - Across cohorts (unit, faculty, time)
 - Are observed differences *real* or the result of chance?
 - What factors are driving the results?
 - What do students consider to be important?
- 

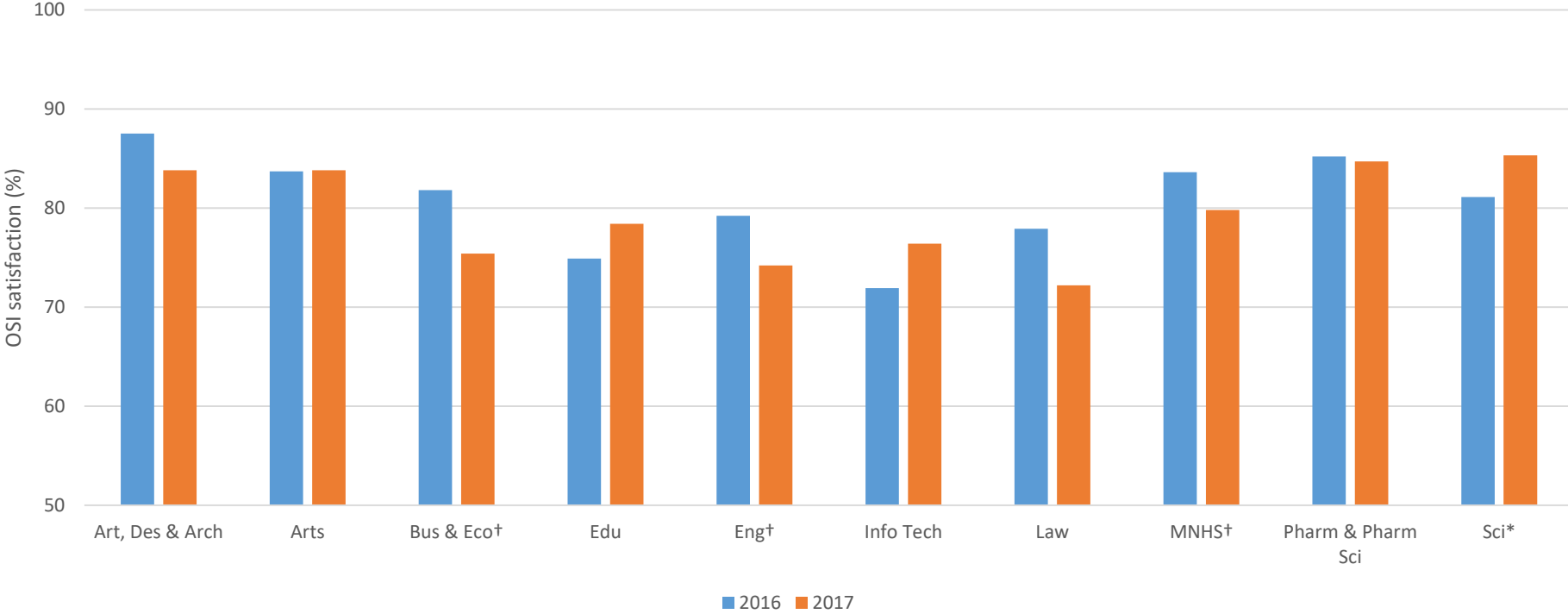
Statistical significance

- Not due only to random chance in the sample
- Function of sample size and effect size
 - Statistical precision increases with the square root of the sample size
- “Yay, the significance level is $p < 0.001!!$ ”
- Statistical vs *practical* significance
- Confidence intervals
 - No overlap = statistically significant
 - Not a replacement for formal hypothesis test
- Finite population correction*

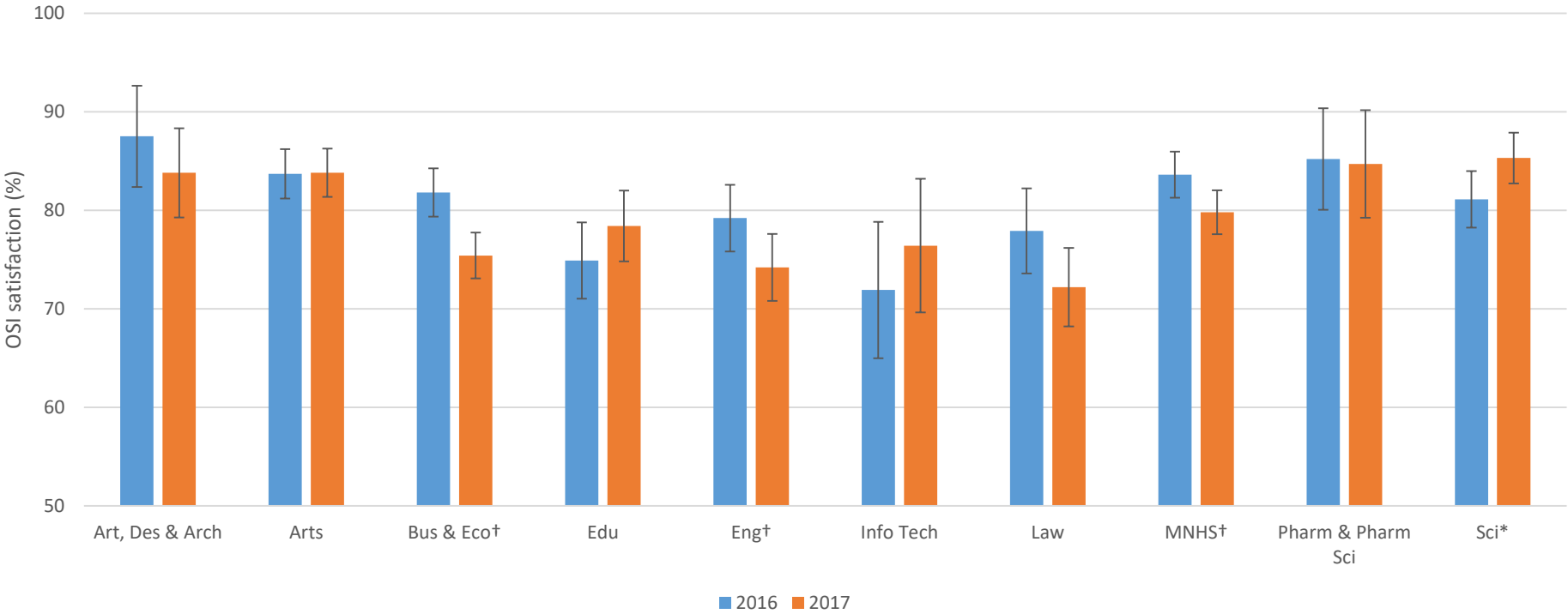
**Statistically
significant is not
necessarily
practically
significant!**

* $FPC = ((N-n)/(N-1))^{1/2}$

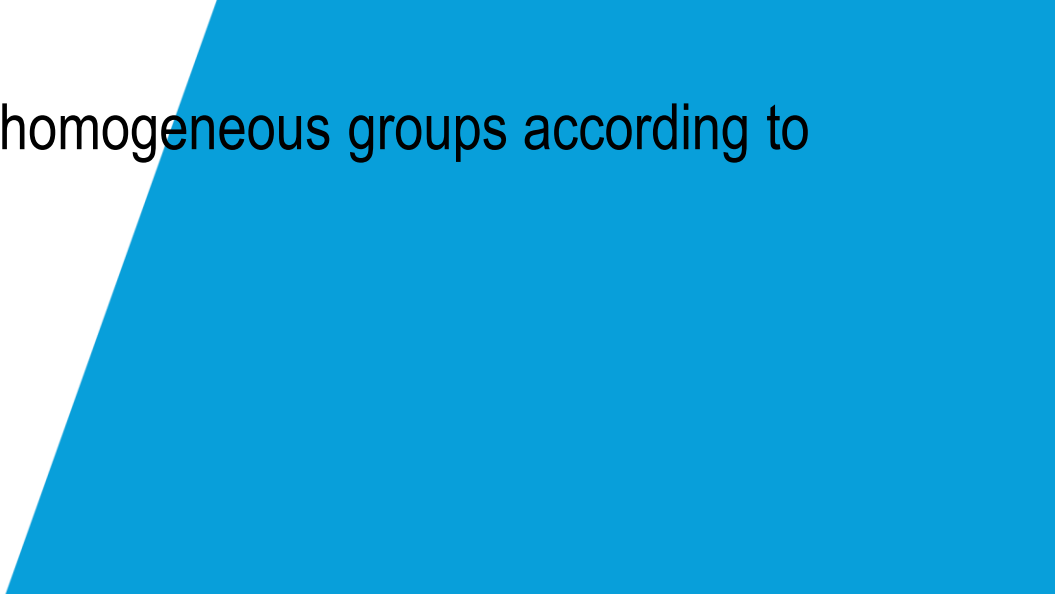
Without confidence intervals...



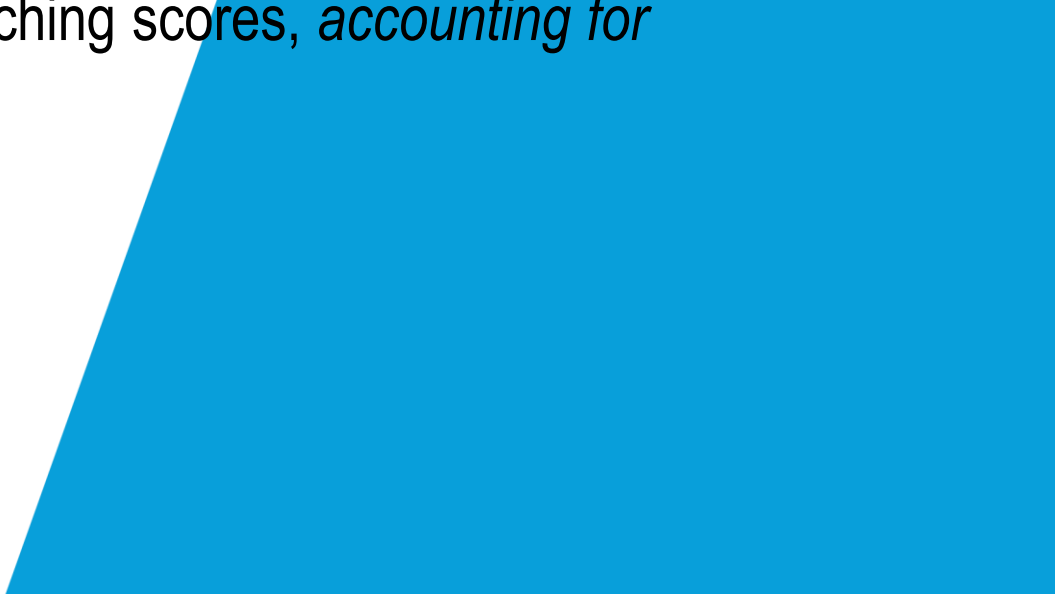
With confidence intervals...




Univariate vs multivariate analysis

- Two approaches to statistical analysis:
 - Univariate: one variable (description)
 - Multivariate: two or more variables (explanation)
 - Dependent and independent variables
 - Confounding variables influence both the DV and IV
 - Spurious correlation
 - Stratification; i.e. dividing your sample into (relatively) homogeneous groups according to a confounding variable
- 

Multivariate research questions

- What characteristics are associated with student satisfaction?
 - Did average satisfaction scores in a unit change significantly between 2017 and 2018, *accounting for changes in student composition?*
 - How do satisfaction scores in a unit compare to other units, *accounting for differences in student composition across units?*
 - What educator characteristics are associated with teaching scores, *accounting for differences in units taught?*
- 

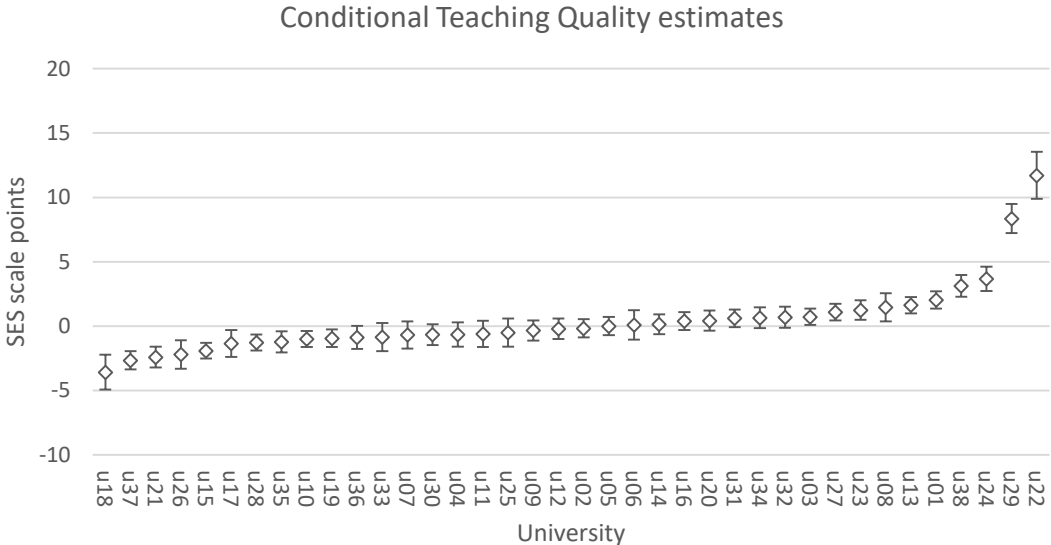
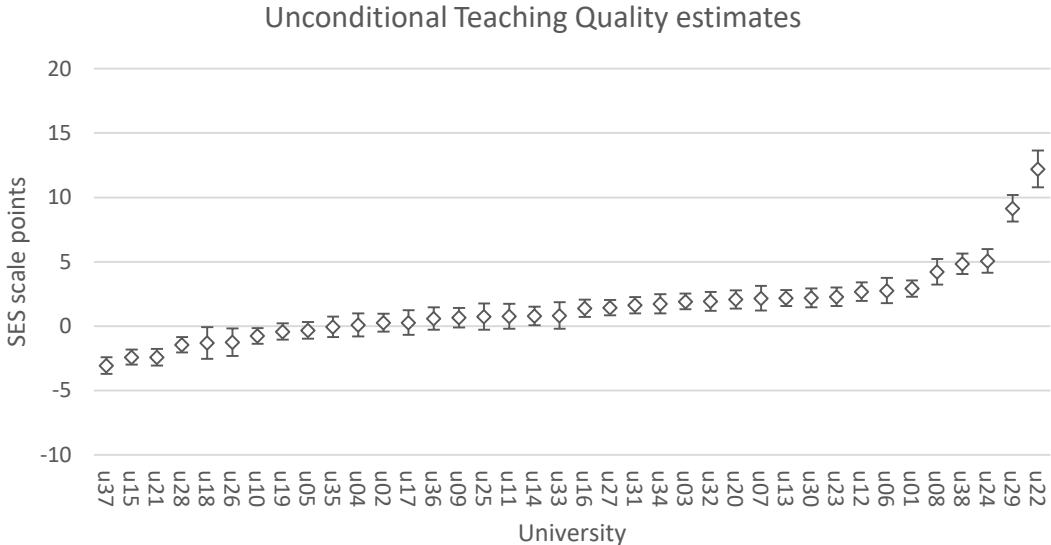
Regression analysis

- Process for estimating the relationship between a dependent variable, and one or more independent variables (or 'predictors')
 - Continuous or quasi-continuous outcome (e.g. psychometric scale): linear regression
 - Binary (yes/no) outcome (e.g. satisfaction indicator): logistic regression
 - Other related techniques:
 - Quantile regression (median)
 - Ordinal regression
 - Multinomial regression
- 

How are variables entered?

- Continuous variables entered as linear effects:
 - “What is the effect on y of a one-unit increase in x ?”
- Categorical variables must be entered as indicator (i.e. dummy) variables:
 - Coded 1 = yes and 0 = no; e.g. sex might be coded as 1 = female and 0 = male
 - “What is the effect of being female (relative to male)?”
 - For categorical variables of >2 categories, include $k-1$ dummy variables, where k is the number of categories.

Student experience by university



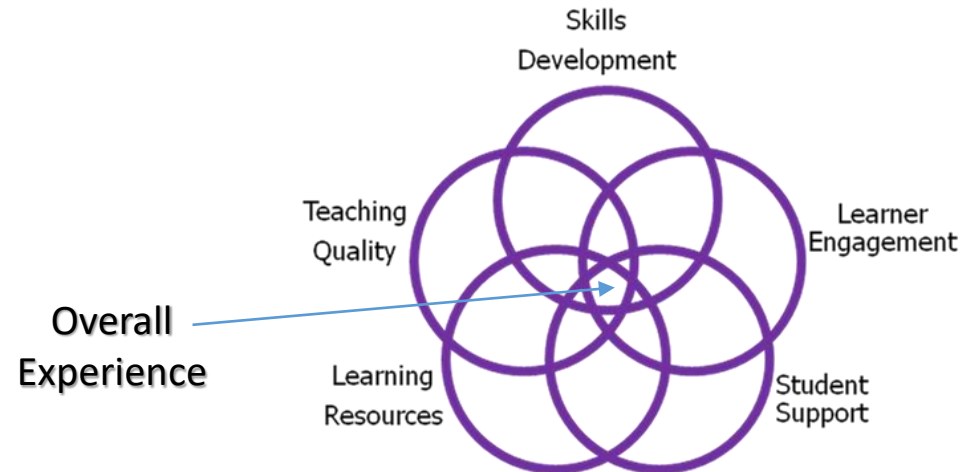
Group	Model (1)	Model (2)
Go8	27	24
ATN	26	22
IRU	20	22
RUN	13	19
Other	16	16

Results on all explanatory variables

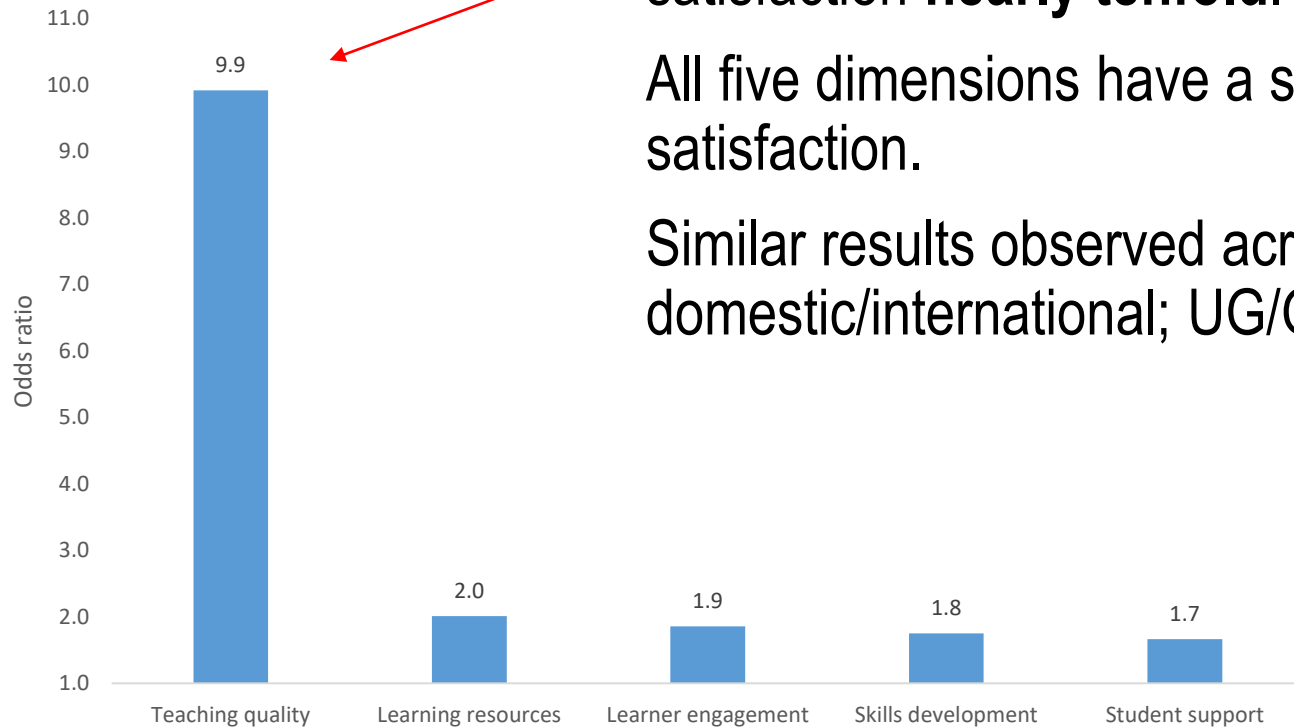
	Model (1)			Model (2)				Model (1)			Model (2)				Model (2)		
Variable	Coef.	s.e.	P	Coef.	s.e.	p	Variable	Coef.	s.e.	P	Coef.	s.e.	p	Variable	Coef.	s.e.	p
University (omitted: go8_8)							run_2	3.715	0.537	0.000	1.459	0.556	0.009	Medicine	4.869	0.512	0.000
go8_1	0.460	0.361	0.203	-0.166	0.362	0.646	run_3*	2.034	0.523	0.000	-0.678	0.539	0.208	Nursing	1.707	0.245	0.000
go8_2	-0.636	0.316	0.044	-0.989	0.315	0.002	run_4	0.339	0.520	0.514	-1.352	0.533	0.011	Pharmacy	2.418	0.580	0.000
go8_3	-2.282	0.311	0.000	-1.909	0.311	0.000	run_5*	0.877	0.513	0.087	-0.599	0.518	0.247	Dentistry	2.867	0.827	0.001
go8_4	0.719	0.392	0.067	-0.328	0.399	0.412	run_6	2.330	0.568	0.000	0.101	0.580	0.862	Vet science	2.649	0.710	0.000
go8_5	2.220	0.323	0.000	1.627	0.324	0.000	oth_1	1.875	0.322	0.000	0.724	0.325	0.026	Rehabilitation	6.271	0.407	0.000
go8_6	-1.350	0.565	0.017	-2.200	0.563	0.000	oth_2	1.709	0.339	0.000	0.613	0.349	0.079	Teaching	2.450	0.243	0.000
go8_7	0.385	0.559	0.490	-0.496	0.555	0.372	oth_3	4.898	0.422	0.000	3.135	0.428	0.000	Humanities	4.860	0.212	0.000
atn_1	-2.567	0.358	0.000	-2.655	0.362	0.000	oth_4*	1.972	0.391	0.000	-0.657	0.411	0.110	Social work	4.658	0.422	0.000
atn_2	1.543	0.321	0.000	1.091	0.326	0.001	oth_5	1.564	0.395	0.000	0.433	0.403	0.283	Psychology	4.738	0.284	0.000
atn_3	1.092	0.386	0.005	0.148	0.388	0.703	oth_6	-0.096	0.344	0.780	-0.940	0.346	0.007	Law	3.860	0.282	0.000
atn_4	1.377	0.357	0.000	0.398	0.359	0.268	oth_7	0.355	0.359	0.323	0.007	0.360	0.984	Creative arts	5.809	0.293	0.000
atn_5	-1.333	0.318	0.000	-1.269	0.320	0.000	oth_8	1.603	0.416	0.000	0.656	0.416	0.115	Communications	4.256	0.322	0.000
iru_1	1.983	0.417	0.000	0.704	0.420	0.094	oth_9	1.911	0.396	0.000	-0.200	0.404	0.620	Tourism	1.735	0.986	0.078
iru_2	-2.562	0.339	0.000	-2.407	0.410	0.000	oth_10	0.024	0.474	0.960	-0.642	0.479	0.180	Other explanatory variables			
iru_3	-1.419	0.679	0.037	-3.574	0.690	0.000	oth_11	12.666	0.940	0.000	11.715	0.936	0.000	Female	0.522	0.118	0.000
iru_4	2.983	0.336	0.000	2.042	0.343	0.000	oth_12	-0.186	0.416	0.654	-1.221	0.421	0.004	Mode: external	1.214	0.286	0.000
iru_5	2.460	0.383	0.000	1.246	0.386	0.001	oth_13*	9.548	0.583	0.000	8.360	0.582	0.000	Mode: mixed	-1.223	0.254	0.000
iru_6	0.995	0.557	0.074	-0.844	0.556	0.129	Field of study (omitted: Business and Mgt)					Attend: part time	0.307	0.220	0.162		
iru_7	0.547	0.453	0.228	-0.860	0.457	0.060	Sci and math				3.725	0.207	0.000	Age: 20-24	0.635	0.141	0.000
run_1	5.103	0.480	0.000	3.671	0.482	0.000	Computing				0.189	0.325	0.561	Age: 25+	2.986	0.179	0.000
Model summary							Engineering				0.152	0.252	0.546	NESB	-1.528	0.149	0.000
Obs	94,350			94,350			Architecture				2.515	0.368	0.000	Overseas student	0.985	0.212	0.000
Prob>F	0.000			0.000			Agriculture				2.577	0.443	0.000	First-in-family student	-0.224	0.109	0.039
R-sq	0.015			0.036			Health services				2.886	0.239	0.000				

What do students consider to be important?

- Useful in terms of optimal allocation of scarce resources
- Explicit importance: asked directly
- Implicit performance: derived statistically
 - Model overall satisfaction (for example) as a function of other survey items
- Derived importance cannot account for differences across individuals



Student experience dimensions underpinning overall satisfaction

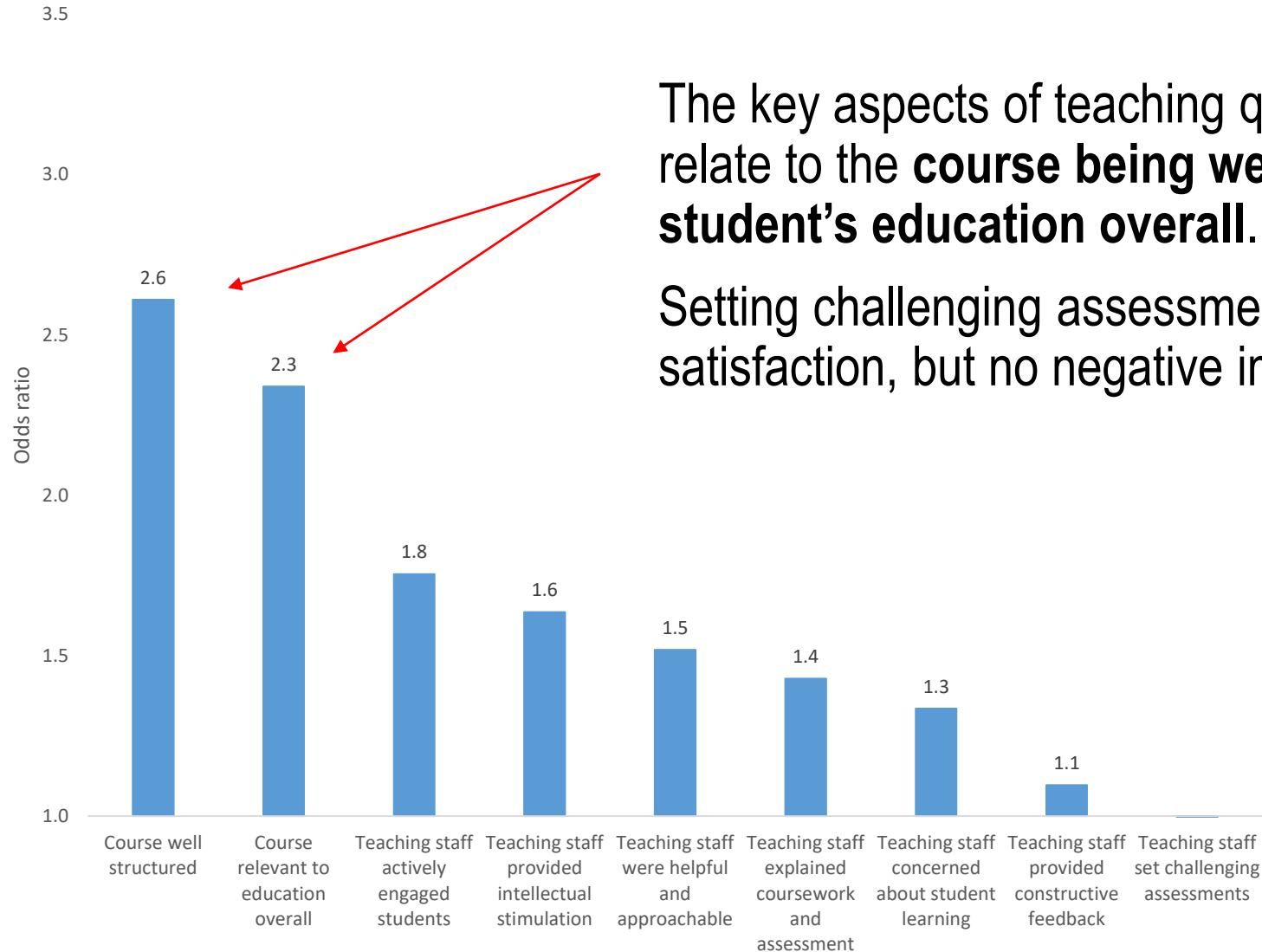


Satisfaction with teaching quality increases the odds of overall satisfaction **nearly tenfold**.

All five dimensions have a significant positive impact on overall satisfaction.

Similar results observed across student cohorts (e.g. domestic/international; UG/GPG).

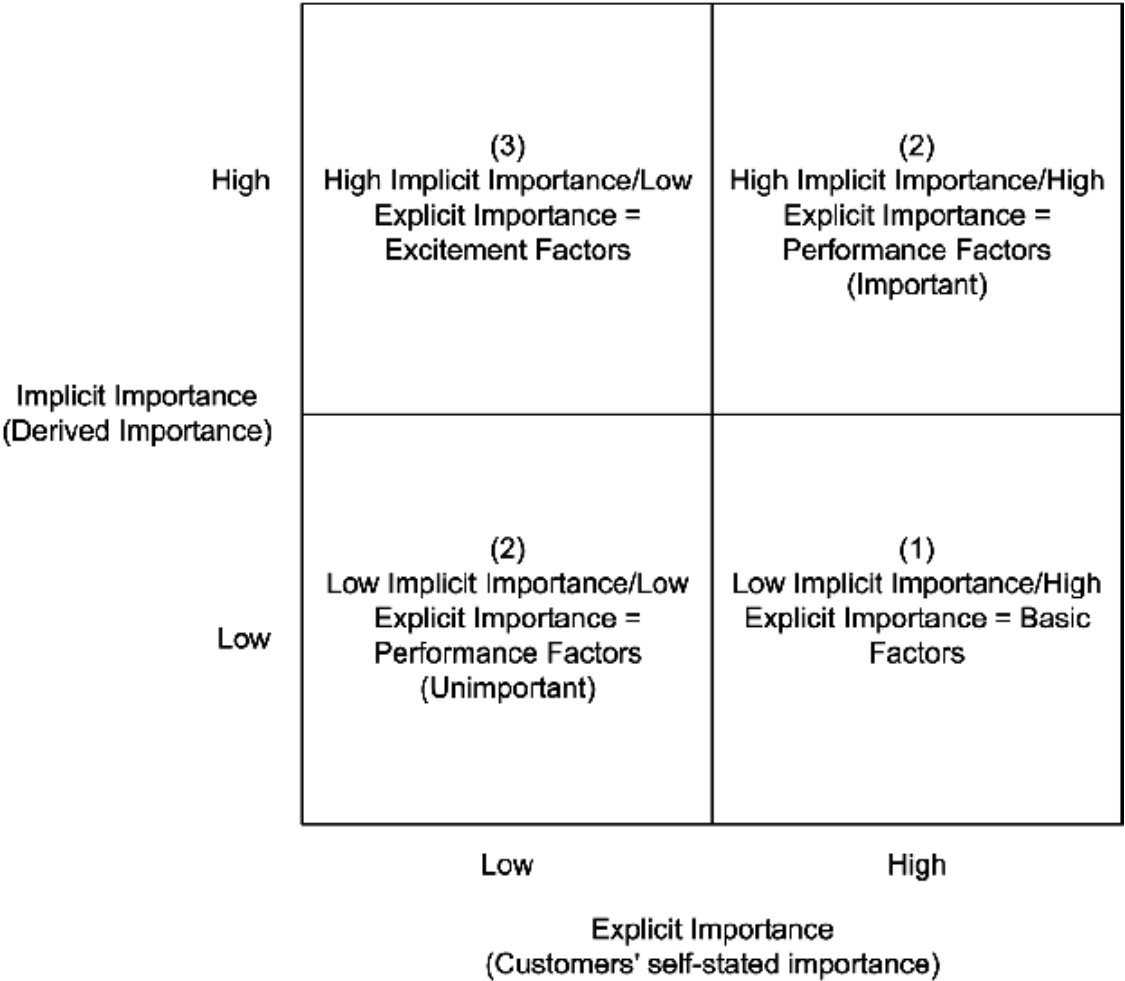
Teaching quality items underpinning overall satisfaction



The key aspects of teaching quality influencing overall satisfaction relate to the **course being well structured** and **relevant to a student's education overall**.


Setting challenging assessments has no positive impact on overall satisfaction, but no negative impact, either.

Explicit vs implicit (derived) importance



Source: Adapted from Vavra, T. G. (1997) Improving your measurement of customer satisfaction: a guide to creating, conducting, analysing, and reporting customer satisfaction measurement programs. Milwaukee, WI: ASQ Quality Press.

Conclusions

- Observed differences are not always real
 - Real differences are not always relevant
 - Beware confounding factors
 - Students place different levels of importance on different facets of their university experience
 - Use statistical methods to derive the most value from your data
- 



Questions?

Dr David Carroll

David.Carroll@monash.edu



MONASH
University