

Research Methods

Statistics with JASP: Scale Development

Andy Gillett

1

Reliability and Validity

Reliability and validity are concepts used to evaluate the quality of the research. They indicate how well a method, technique, or test measures something. Reliability is about the consistency of a measure, and validity is about the accuracy of a measure.

Reliability

The extent to which a measure produces consistent, dependable, and repeatable results under the same conditions.

- Test-Retest Reliability - evaluates how consistent a measure's results are over time.
- Inter-Rater Reliability - measures how consistently two or more independent observers agree when rating, coding, or assessing the same phenomenon.
- Internal Consistency - is the consistency of the measurement itself. Do you get the same results from different parts of a scale that are designed to measure the same thing?

Validity

The extent to which a measure accurately reflects the concept or construct it's intended to measure.

- Content Validity - measures the appropriateness of the content and composition of the scale, given what is being measured?
- Construct Validity - measures the underlying construct that the scale is designed to measure.
- Criterion Validity - measures whether or not the results obtained from the scale can be used to predict something related to the scale.

2

Student Satisfaction

- Teaching Quality
- Course Content
- Support and Guidance
- Rooms and Facilities
- Resources

- 6 Extremely satisfied
- 5 Very satisfied
- 4 Moderately satisfied
- 3 Moderately dissatisfied
- 2 Very dissatisfied
- 1 Extremely dissatisfied

3

Student Satisfaction

Teaching	Teaching	Teaching	Teaching	Content 1	Content 2	Content 3	Content 4	Content 5	Support 1	Support 2	Support 3	Support 4	Facilities 1	Facilities 2	Facilities 3	Facilities 4
1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5
2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1
4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2
5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3
6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4
1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5
2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1
4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2
5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3
6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4
1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5
2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1
4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2
5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3
6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4
1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5
2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1
4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2
5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3
6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4
1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5
2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1
4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2
5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3
6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4
1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5
2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1
4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2
5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3
6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4

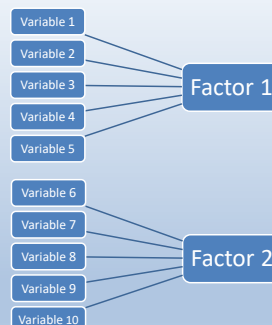
4

Factor Analysis

- Variable 1
- Variable 2
- Variable 3
- Variable 4
- Variable 5
- Variable 6
- Variable 7
- Variable 8
- Variable 9
- Variable 10

5

Factor Analysis



Eigenvalue - The amount of the total variance explained by that factor is called the **eigenvalue** of a factor.

Factor Loading - The extent to which a particular variable contributes to a given factor is the **factor loading** of a variable.

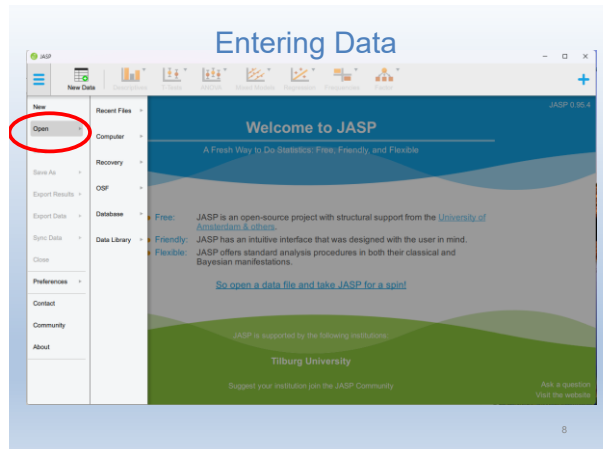
6

5

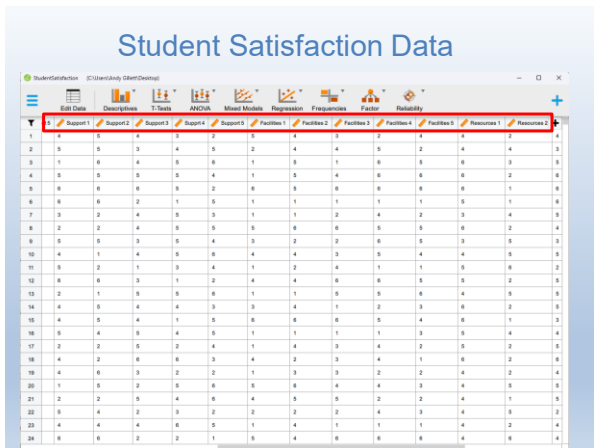
6



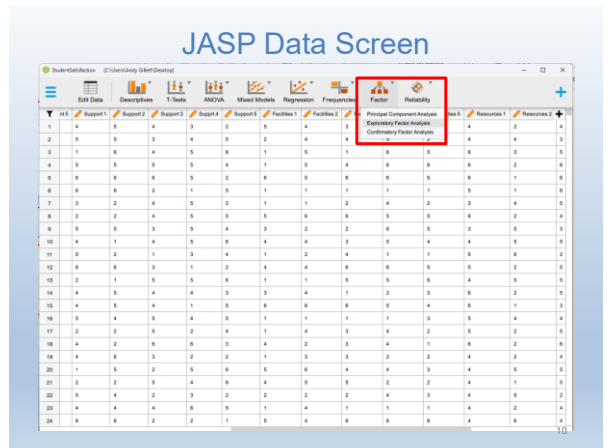
7



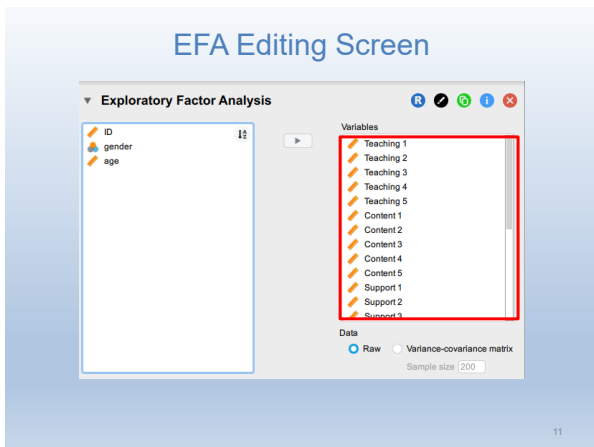
8



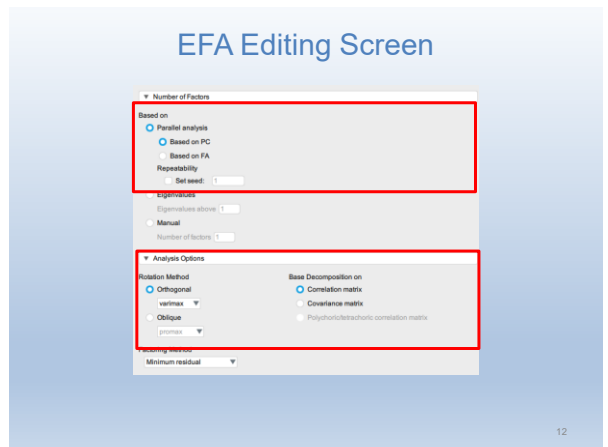
9



10



11



12

EFA Editing Screen

Output Options

Display loadings above

0.4

Tables

- Structure matrix
- Factor correlations
- Additional R indices
- Residual matrix
- Parallel analysis
- Based on PC
- Based on FA

Assumption checks

- KMO test
- Bartlett's test
- Mardia's test
- Anti-image correlation matrix
- Add FA scores to data

Order Loadings By

- Size
- Variables

Plots

- Path diagram
- Scree plot
- Parallel analysis results

Missing Values

- Exclude cases pairwise
- Exclude cases listwise

Probs: FA

13

EFA Results

Kaiser-Meyer-Olkin (KMO) 0.807

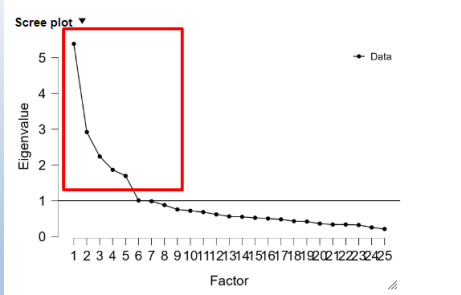
MSA	
Overall MSA	0.808
Content 1	0.888
Content 2	0.887
Content 3	0.748
Content 4	0.793
Content 5	0.842
Facilities 1	0.738
Facilities 2	0.712
Facilities 3	0.788
Facilities 4	0.825
Facilities 5	0.854
Resources 1	0.848
Resources 2	0.888
Resources 3	0.818
Resources 4	0.743
Resources 5	0.780
Support 1	0.828
Support 2	0.845
Support 3	0.814
Support 5	0.918
Support 4	0.889
Teaching 1	0.884
Teaching 2	0.838
Teaching 3	0.829
Teaching 4	0.835
Teaching 5	0.884

Bartlett's Test

X ²	df	p
24895.550	300.000	< .001

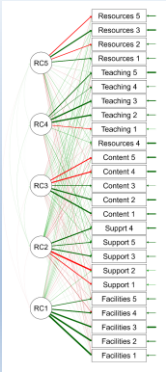
14

EFA Result



15

EFA Result



16

JASP Editing Screen

Factor Loadings

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Uniqueness
Facilities 1	0.776					0.347
Facilities 2	0.761					0.399
Facilities 3	0.714					0.474
Facilities 4	0.807					0.449
Facilities 5	0.518					0.632
Support 2		-0.725				0.353
Support 4		0.631				0.425
Support 3		0.602				0.478
Support 1		-0.536				0.696
Support 5		0.521				0.575
Content 1			0.683			0.480
Content 2			0.856			0.524
Content 4			-0.628			0.475
Content 5			-0.585			0.519
Content 3			0.483			0.747
Teaching 2				0.693		0.480
Teaching 3				0.690		0.442
Teaching 5				0.549		0.531
Teaching 1				-0.473		0.719
Teaching 4				0.424		0.665
Resources 3					0.647	0.414
Resources 1					0.534	0.626
Resources 5					-0.531	0.679
Resources 2					-0.504	0.682
Resources 4						0.746

Note: Applied rotation method is varimax.

17

EFA Results

Factor Characteristics

	Unrotated solution			Rotated solution		
	Eigenvalue	Proportion var.	Cumulative	SumSq. Loadings	Proportion var.	Cumulative
Factor 1	4.890	0.196	0.196	2.756	0.110	0.110
Factor 2	2.435	0.097	0.293	2.654	0.106	0.216
Factor 3	1.680	0.067	0.360	2.243	0.090	0.306
Factor 4	1.282	0.051	0.411	2.145	0.086	0.392
Factor 5	1.155	0.046	0.458	1.644	0.066	0.458

18

13

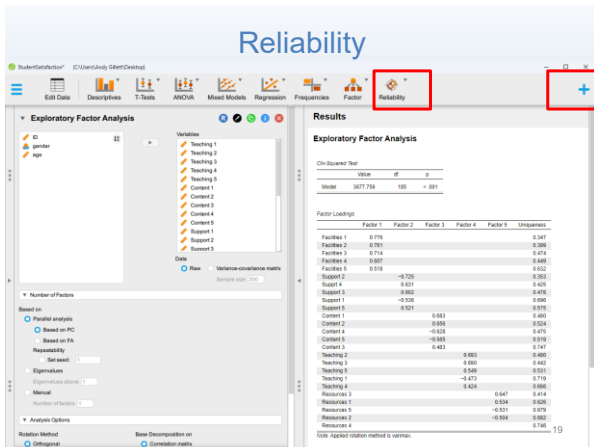
14

15

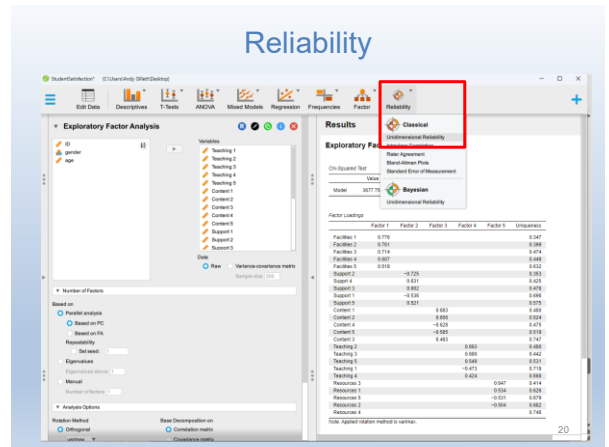
16

17

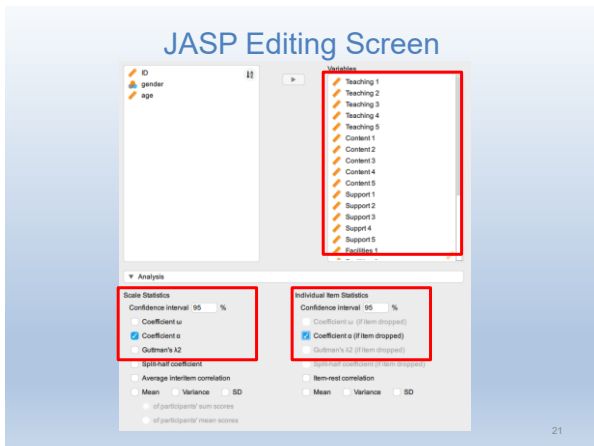
18



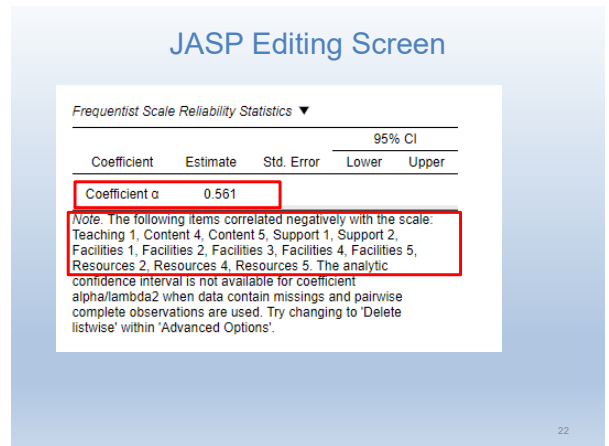
19



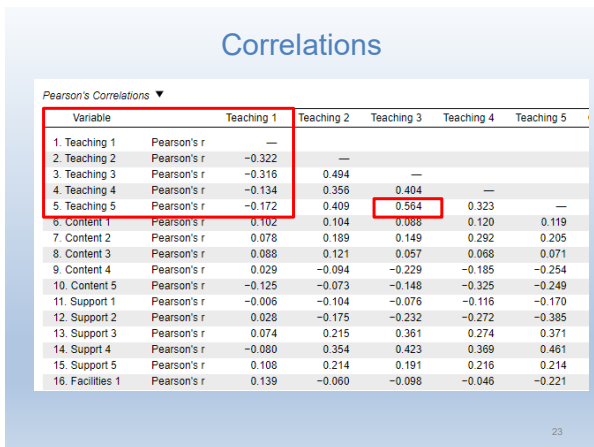
20



21



22



23



24

Reverse-Scaled Items

▼ Reverse-Scaled Items

Normal-Scaled Items

- Teaching 2
- Teaching 3
- Teaching 4
- Teaching 5
- Content 1
- Content 2
- Content 3
- Support 3
- Support 4
- Support 5
- Resources 1
- Resources 3

Reverse-Scaled Items

- Teaching 1
- Content 4
- Content 5
- Support 1
- Support 2
- Facilities 1
- Facilities 2
- Facilities 3
- Facilities 4
- Facilities 5
- Resources 2
- Resources 4
- Resources 5

25

Internal Consistency: Results

Frequentist Scale Reliability Statistics ▼

Coefficient	Estimate	Std. Error	95% CI	
			Lower	Upper
Coefficient α	0.829			

Note. The analytic confidence interval is not available for coefficient alpha/lambda2 when data contain missings and pairwise complete observations are used. Try changing to 'Delete listwise' within 'Advanced Options'.

26

Internal Consistency: Results

Frequentist Scale Reliability Statistics ▼

Coefficient	Estimate	Std. Error	95% CI	
			Lower	Upper
Coefficient α	0.829			

Note. The analytic confidence interval is not available for coefficient alpha/lambda2 when data contain missings and pairwise complete observations are used. Try changing to 'Delete listwise' within 'Advanced Options'.

Frequentist Individual Item Reliability Statistics

Item	Estimate	Lower 95% CI	Upper 95% CI
Teaching 1	0.834		
Teaching 2	0.826		
Teaching 3	0.822		
Teaching 4	0.823		
Teaching 5	0.818		
Content 1	0.822		
Content 2	0.826		
Content 3	0.831		
Content 4	0.816		
Content 5	0.817		
Support 1	0.824		
Support 2	0.811		
Support 3	0.821		
Support 4	0.818		
Support 5	0.820		
Facilities 1	0.821		
Facilities 2	0.825		
Facilities 3	0.821		
Facilities 4	0.817		
Facilities 5	0.827		
Resources 1	0.825		
Resources 2	0.829		
Resources 3	0.821		
Resources 4	0.834		
Resources 5	0.827		

27

Internal Consistency: Results

Frequentist Scale Reliability Statistics ▼

Coefficient	Estimate	Std. Error	95% CI	
			Lower	Upper
Coefficient α	0.829			

Note. The analytic confidence interval is not available for coefficient alpha/lambda2 when data contain missings and pairwise complete observations are used. Try changing to 'Delete listwise' within 'Advanced Options'.

Content 1	0.822
Content 2	0.826
Content 3	0.831
Content 4	0.816
Content 5	0.817

28

Internal Consistency: Results

▼ Unidimensional Reliability

Teaching 4

Teaching 5

Content 1

Content 2

Content 3

Content 4

Content 5

Support 1

Support 2

Support 3

Support 4

Support 5

Resources 1

Resources 2

Resources 3

Resources 4

Resources 5

Facilities 1

Facilities 2

Facilities 3

Facilities 4

Facilities 5

29

JASP Editing Screen

Frequentist Scale Reliability Statistics ▼

Coefficient	Estimate	Std. Error	95% CI	
			Lower	Upper
Coefficient α	0.811			

Note. The analytic confidence interval is not available for coefficient alpha/lambda2 when data contain missings and pairwise complete observations are used. Try changing to 'Delete listwise' within 'Advanced Options'.

Frequentist Individual Item Reliability Statistics

Item	Estimate	Lower 95% CI	Upper 95% CI
Facilities 1	0.751		
Facilities 2	0.765		
Facilities 3	0.747		
Facilities 4	0.790		
Facilities 5	0.813		

Note. The analytic confidence interval is not available for coefficient alpha/lambda2 when data contain missings and pairwise complete observations are used. Try changing to 'Delete listwise' within 'Advanced Options'.

30

Internal Consistency: Results

Low Cronbach's alpha

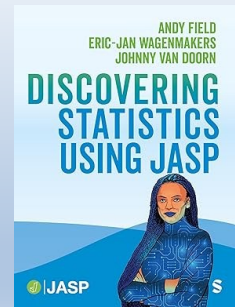
Maybe the underlying model is wrong.
Cronbach's Alpha assumes unidimensionality. That could be wrong in your case.

What does EFA tell you?

31

31

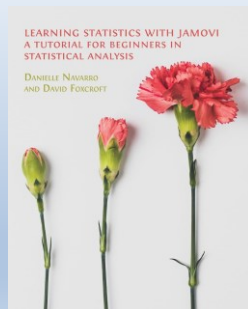
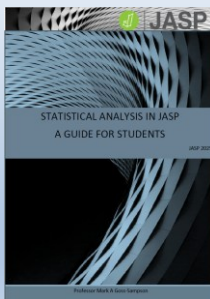
Reading



32

32

Reading



33

33

References & Further Reading

- Brown, J. D. & Rogers, T. S. (2002). *Doing second language research*. Oxford: Oxford University Press.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. New York: Routledge.
- Dörnyei, Z. (2007). *Research methods in applied linguistics*. Oxford: Oxford University Press.
- De Winter, J. C. F. & Dodou, D. (2012). Five point Likert items: t-test versus Mann-Whitney-Wilcoxon. *Practical Assessment, Research and Evaluation*, 15(11).
- DeVellis, R. F. (1991). *Scale development: Theory and applications*. London: Sage.
- Field, A., van Doorn, J. & Wagenmakers, E.-J. (2025). *Discovering statistics using JASP*. London: Sage.
- Hand, D. J. (2008) *Statistics: A very short introduction*. Oxford: Oxford University Press.
- Goss-Sampson, M. A. (2025). *Statistical analysis in JASP: A guide for students*. JASP.
- Halter, C. P. (2024). *Quantitative Analysis with JASP open-source software: Foundational Methods*. Author.
- Hatch, E. & Lazaraton, A. (1991). *The research manual: Design and statistics for applied linguistics*. Boston, MA: Heinle & Heinle.
- Johnson, R. L. & Morgan, G. B. (2016). *Survey scales: A guide to development, analysis and reporting*. London: The Guilford Press.
- Lowie, W. & Seton, B. (2013). *Essential statistics for applied linguistics*. London: Palgrave Macmillan.
- Navarro, D.J., Foxcroft, D.R., & Faulkenberry, T.J. (2025). *Learning Statistics with JASP: A Tutorial for Psychology Students and Other Beginners*. (Version 2.0). <https://learnstatswithjasp.com>
- Pett, M. A., Lackey, N. R. & Sullivan, J. J. (2003). *Making sense of factor analysis*. London: Sage.
- Popham, W. J. , & Sirotnik, K. A. (1973). *Educational statistics: Use and interpretation* (2nd ed.). New York: Harper & Row
- Salkind, N. J. (2011). *Statistics for people who (think they) hate statistics* (4th ed.). London: Sage.
- Saunders, M. & Lewis, P. (2012). *Doing research in business and management*. London: FT/Prentice Hall.
- Saunders, M., Lewis, P. & Thornhill, A. (2016). *Research methods for business students* (7th ed.). London: FT/Prentice Hall.
- Shavelson, R. (1988). *Statistical reasoning for the behavioral sciences*. Boston: Allyn and Bacon.
- Spiegelhalter, D. (2020). *The art of statistics: Learning from data*. London: Pelican.

34

34